

THESIS / THÈSE

MASTER EN SCIENCES INFORMATIQUES

Visualisation d'informations boursières sur téléphones mobiles

Custinne, Grégory

Award date:
2004

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

MÉMOIRE

Visualisation d'informations boursières

sur téléphones mobiles

présenté par

GRÉGORY CUSTINNE

Mémoire en vue d'obtenir le grade de

Maître en Informatique

Abstract

This dissertation present new methods for visualization of stock market data on mobile phones. Because of the small height of mobile phone's screen, its weak resolution and the great volume of data to represent, we suggest to condense beforehand the information. Data are initially treated with a standard technic of classification before being visualized. The classification used is the not-supervised classification or clustering. This one has lots of technics but that which we choose is the K-Means technics or the mobile centers technics. For the visualization, we propose two methods of standard representation (Bar Chart et Tableau) and two more recent methods (Pixel Bar Chart et Spirale) which present more informations to users than the two others representations.

Key-words : visualization, mobile phone, classification, K-Means.

Résumé

Ce mémoire présente de nouvelles méthodes pour la visualisation de données boursières sur téléphones mobiles. Vu la taille très réduite de l'écran d'un téléphone mobile, sa faible résolution et le grand volume de données à représenter, nous suggérons de condenser préalablement l'information. Les données sont d'abord traitées avec une technique standard de classification avant d'être visualisées. La classification utilisée est la classification non-supervisée ou clustering. Celle-ci possède plusieurs techniques mais celle que nous avons choisi est la technique des K-Means ou des centres mobiles. Pour la visualisation, nous proposons deux méthodes de représentation classique (Bar Chart et Tableau) et deux méthodes plus récentes (Pixel Bar Chart et Spirale) qui présentent plus d'informations aux utilisateurs que les deux autres représentations.

Mots-clés : visualisation, téléphone mobile, classification, K-Means.

Je voudrais tout d'abord commencer par remercier mon professeur et promoteur, Mrs Noirhomme, pour toute l'aide apportée tout au long de ce travail et pour m'avoir permis d'écrire un article sur le même sujet que ce mémoire.

Ensuite, j'aimerais aussi remercier mon superviseur à Udine, Mr Chittaro, également pour son aide lors de mon séjour en Italie, et surtout pour avoir orienter mon travail.

Je terminerais finalement en remerciant toutes les personnes du département HCI à Udine pour leur accueil chaleureux et à ma famille pour son support indirect mais nécessaire.

Table des matières

1	Introduction	9
2	État de l'art	11
2.1	Techniques de visualisation de l'information	11
2.1.1	Classification des types de représentations	11
2.2	Visualisations boursières pour téléphones mobiles sur Sites Wap	15
2.2.1	Les tableaux	16
2.2.2	Les graphiques	19
2.3	Application existante	20
2.3.1	Midcast	20
2.4	Conclusion	23
3	Analyse des besoins	24
3.1	Utilisateurs	25
3.1.1	Types d'utilisateur	25
3.1.2	Les caractéristiques de l'utilisateur	25
3.1.3	Environnement	27
3.1.4	Tâches	28
4	Limites des téléphones mobiles	30
4.1	Limitation par rapport à l'écran	30
4.2	Limitation par rapport à la mémoire	31
4.3	Limitation par rapport à la vitesse du processeur	31
4.4	Problème de navigation	32
4.5	Conclusion	32
5	Classification	33
5.1	La méthode des K-Means	34
5.2	Exemple de calcul des K-Means	35
5.3	Apport de cette méthode	35
5.4	Conclusion	36

6	Les visualisations boursières de l'application	37
6.1	Introduction	37
6.2	Visualisations Classiques	37
6.2.1	Le Bar Chart	37
6.2.2	Le tableau	40
6.3	Les nouvelles visualisations	41
6.3.1	Le pixel Bar Chart	41
6.3.2	La spirale [Weber]	45
6.4	Conclusion	47
7	Implémentation	48
7.1	Choix du langage	48
7.1.1	Introduction	48
7.1.2	Java Micro Edition	48
7.2	Démarche	53
7.2.1	Introduction	53
7.2.2	Architecture de l'application	53
7.2.3	Récupération des données boursières	54
7.2.4	Le calcul des K-Means	54
7.2.5	Implémentation des différentes visualisations	56
7.3	Interface	57
7.4	Problèmes et Solutions	59
7.4.1	Lenteur lors de la récupération des données	59
7.4.2	Lenteur lors de l'affichage de certaines visualisations	60
8	Expérience	61
8.1	Introduction	61
8.2	Evaluation	61
8.2.1	L'échantillon	61
8.2.2	Présentation du questionnaire	62
8.2.3	Utilisabilité	64
8.2.4	Résultats	64
8.3	Améliorations	78
8.3.1	La navigation	78
8.3.2	Les différentes visualisations	79
8.3.3	Opportunité : le son	79
9	Conclusions	81

A	Appendice	85
A.1	Les différents questionnaires	85
A.1.1	Utilisateur 1	85
A.1.2	Utilisateur 2	86
A.1.3	Utilisateur 3	86
A.1.4	Utilisateur 4	87
A.1.5	Utilisateur 5	87
A.1.6	Utilisateur 6	88
A.1.7	Utilisateur 7	89
A.1.8	Utilisateur 8	89
A.1.9	Utilisateur 9	90
A.1.10	Utilisateur 10	90
A.1.11	Utilisateur 11	91
A.1.12	Utilisateur 12	92
A.1.13	Utilisateur 13	93
A.1.14	Utilisateur 14	93
A.1.15	Utilisateur 15	94
A.1.16	Utilisateur 16	95
A.1.17	Utilisateur 17	95
A.1.18	Utilisateur 18	96
A.1.19	Utilisateur 19	97
A.1.20	Utilisateur 20	97
A.2	Table des valeurs critiques pour la t distribution de Student	99

Table des figures

2.1	Graphique à secteur	12
2.2	Pyramide des âges	12
2.3	Graphique linéaire simple	13
2.4	Graphique linéaire multiple	13
2.5	Nuage de points	13
2.6	Nuages de points avec aide à la visualisation	14
2.7	Diagramme à bâtons ou Bar Chart	15
2.8	Histogramme	15
2.9	Menu principal (www.wap.boursorama.com)	17
2.10	Carte 1, choix de l'indice. Correspond au choix d'une ligne dans le tableau.	17
2.11	Carte 2, lien détail vers le reste des informations de l'indice.	17
2.12	Carte 3, ligne du tableau.	18
2.13	Choix de l'indice.	18
2.14	Informations sur l'indice.	18
2.15	Cours de l'action SONY lors de la dernière heure.	19
2.16	Historique du cours de l'action SONY lors des dernières heures.	19
2.17	Cour de l'action SIE sur le site de http://wap.moneyextra.com	20
2.18	Cour d'une action sur le site de www.firsinvest.waptoo.com	20
2.19	Tableau comprenant les actions du portefeuille d'un utilisateur.	21
2.20	Candlestick d'une action.	22
2.21	Graphe d'une action lors des trois dernières minutes.	22
2.22	Informations boursières d'une action.	22
5.1	Exemple du principe des K-Means où $K=2$	34
6.1	Bar Chart de type 1	38
6.2	Bar Chart de type 1 sur téléphone mobile	38
6.3	Bar Chart de type 2	39
6.4	Bar Chart de type 2 sur téléphone mobile	39
6.5	Bar Chart de type 1 après avoir fait défiler l'écran vers le bas pour consulter le contenu du cluster 1	40

6.6	Bar Chart de type 1 avec contenu du cluster 4	40
6.7	Représentation d'un tableau dans l'application	41
6.8	Principe du Pixel Bar Chart [Kei]	42
6.9	Bar Chart avec des barres de même largeur. Ce Bar Chart représente le nombre de clients pour 12 type de produits différents[Keim]	42
6.10	Bar Chart avec des barres de même hauteur. La largeur d'un rectangle représente ici le nombre de client pour le type de produit correspondant[Keim]	43
6.11	Pixel Bar Chart avec des barres de même largeur. L'attribut séparant est toujours le nombre de clients par type de produit. Le nombre de visite et la quantité d'argent donnée pour le produit sont les attributs x et y[Keim]	43
6.12	Pixel Bar Chart avec des barres de même largeur. Tout comme la figure précédente, l'attribut séparant est le nombre de clients par type et les attributs x et y sont le nombre de visite et la quantité d'argent donnée pour le produit[Keim]	44
6.13	Pixel Bar Chart sur téléphone mobile	44
6.14	Pixel Bar Chart sur téléphone mobile avec contenu du cluster 2	45
6.15	Pixel Bar Chart sur téléphone mobile avec contenu du cluster 7	45
6.16	Spirale	46
6.17	Visualisation de la spirale sur téléphone mobile	46
7.1	Java 2 plate-forme	49
7.2	CDC et CLDC	50
7.3	Les packages CLDC et MIDP	52
7.4	Les Midlets	53
7.5	Menu principal de l'application	57
7.6	Choix de la période (choix exclusif).	57
7.7	Choix du domaine des valeurs (choix exclusif).	58
7.8	Choix du nombre de clusters.	58
7.9	Menu des visualisations.	58
7.10	Choix du premier cluster représentant la spirale extérieure. Le choix du deuxième suit celui-ci et est le même.	59
7.11	Texte apparaissant à l'écran pendant quelques secondes si le numéro du cluster choisi est supérieur au nombre de clusters choisi par l'utilisateur au préalable.	59

Liste des tableaux

3.1	Evolution du nombre d'abonnés mobiles dans le monde	24
4.1	Tailles des écrans de certains Nokia.	30
4.2	Capacité mémoire de certains Nokia.	31
7.1	Liste des packages de CLDC	51
7.2	Liste des packages de MIDP	51
8.1	Informations personnelles.	62
8.2	Expérience.	62
8.3	Tableau des différentes valeurs qualitatives.	64
8.4	Tableau des fréquences initiales observées.	65
8.5	Tableau des fréquences observées.	65
8.6	Tableau des fréquences théoriques.	65
8.7	Calcul du χ^2_{obs}	66
8.8	Tableau des effectifs et des effectifs cumulés du Bar Chart de type 1.	67
8.9	Tableau des effectifs et des effectifs cumulés du Bar Chart de type 2.	67
8.10	Tableau des effectifs et des effectifs cumulés du Tableau.	68
8.11	Tableau des effectifs et des effectifs cumulés du Pixel Bar Chart.	68
8.12	Tableau des effectifs et des effectifs cumulés de la Spirale.	68
8.13	Tableau des valeurs ordinales et des préférences correspondantes.	69

Chapitre 1

Introduction

Une énorme masse d'information est maintenant disponible en temps réel grâce à l'usage d'Internet. Cet accès permanent facilite et accélère la prise de décision. C'est particulièrement le cas dans le domaine boursier où les courtiers, ou d'autres utilisateurs, recherchent un maximum d'information sur les marchés en un minimum de temps. L'usage des téléphones mobiles facilite cette tâche car cette technologie permet d'avoir accès à l'information non plus seulement à partir d'un ordinateur localisé dans un bureau mais à partir de n'importe quel endroit.

Ce mémoire a pour but de présenter une approche basée sur la combinaison de deux méthodes : la classification non supervisée, ou clustering, et la visualisation. La première permet de résumer l'information à fournir, la deuxième donne une perception globale d'une grande masse de données. Deux méthodes de visualisation classiques (Bar Chart et Tableau) pour représenter des valeurs moyennes ou différentielles sur une période de temps et deux méthodes récentes (Pixel Bar Chart et Spirale) pour représenter l'évolution d'une valeur au cours du temps vont être présentées. La combinaison de ces méthodes permettra d'afficher sur un petit écran un grand volume de données.

Le cas particulier traité ici concerne les données boursières. Le système proposé est destiné à aider toute personne s'intéressant au domaine boursier ou ayant un grand portefeuille d'actions. Grâce à ce système, son utilisateur pourra visualiser plus d'actions sur un téléphone mobile c'est-à-dire consulter beaucoup plus d'actions à la fois partout où il est. De plus, il aura à sa disposition d'une part des visualisations faciles à utiliser mais ne donnant guère d'informations et d'autre part des nouvelles visualisations plus sophistiquées qui, avec un apprentissage préalable, pourront lui apporter beaucoup plus d'informations en un coup d'oeil.

Dans le chapitre 2, nous allons nous intéresser aux différentes méthodes de visualisations d'information. Nous allons aussi observer celles disponibles sur les sites Wap et nous finirons par jeter un coup d'oeil à une application existante : Midcast.

Le chapitre 3 traitera d'une analyse des besoins pour cette application.

Dans le chapitre 4, nous allons présenter les différentes limites des téléphones mobiles. Il est nécessaire de les présenter pour bien comprendre le but de la combinaison des méthodes que sont le clustering et la visualisation.

Dans le chapitre 5, nous étudierons le principe de la classification et plus spécialement la technique des centres mobiles ou K-Means qui est une des deux méthodes utilisées par l'application. Nous verrons aussi l'apport de cette méthode aux téléphones portables.

Le chapitre 6 sera consacrée aux différentes visualisations de l'application à savoir le Bar Chart, le Tableau, le Pixel Bar Chart et la Spirale.

Dans le chapitre 7, nous regarderons de plus près la structure de l'application ainsi que l'implémentation de la technique des K-Means et des différentes visualisations. Des problèmes de l'application seront aussi présentés ainsi que des solutions possibles.

Dans le chapitre 8, nous essayerons d'évaluer l'application en présentant d'abord l'interface puis un questionnaire d'évaluation. Les réponses des utilisateurs et analyses de celles-ci seront présentées par la suite.

Le mémoire se terminera par une conclusion avec des améliorations possibles dans le chapitre 10.

Chapitre 2

État de l'art

2.1 Techniques de visualisation de l'information

2.1.1 Classification des types de représentations

Il existe différents types de visualisation générale, nous pouvons en distinguer 6 catégories principales classées de la représentation la plus proche du réel (analogie au monde réel) à la représentation la plus abstraite (qui fait appel à l'arbitraire).

La photographie

La photographie est la représentation la plus proche du réel mais n'en est pas une reproduction exacte, son interprétation diffère d'utilisateur en utilisateur.

Les schémas

Cette catégorie regroupe les dessins, les croquis, les schémas. Le degré de ressemblance avec l'objet schématisé dépend de la qualité du schéma utilisé. Un dessin détaillé peut être fortement ressemblant au réel contrairement à un schéma qui par sa simplicité peut demander un plus gros effort de compréhension de la part du lecteur. Un dessin sera souvent associé à un objet réel, tandis qu'un schéma pourra être utilisé pour des principes plus abstraits.

les graphiques

Les graphiques sont utilisés pour représenter des données quantitatives ou nominales, il existe par ailleurs de nombreux types de graphiques différents.

1. Graphiques à barre : Voici quelques exemples de graphiques à barres : Graphiques à barres verticales, horizontales, à barres empilées, par points.

2. Graphiques figuratifs : Ce sont de simples représentations dans lesquelles on introduit des symboles. Exemple :

* = 2 bonbons

Luc : **

Greg : ****

3. Graphiques à secteur (fromages) : figure 2.1.

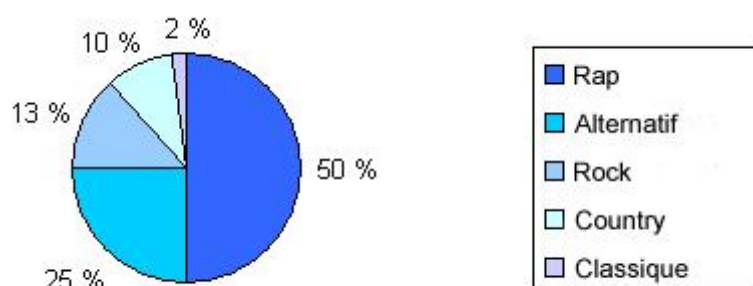


FIG. 2.1 – Graphique à secteur

Dans certains cas il vaut mieux utiliser les graphiques à barres que les fromages, notamment pour comparer deux graphiques. Exemple : la pyramide des âges.

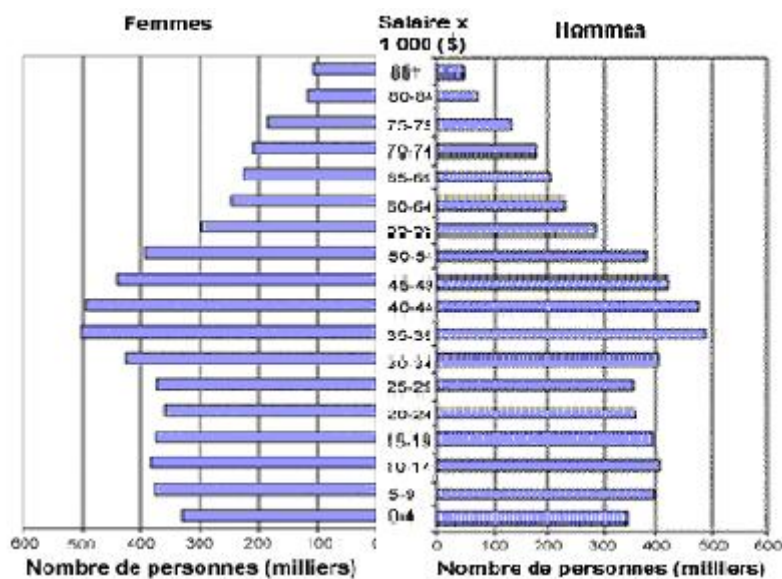


FIG. 2.2 – Pyramide des âges

4. Graphiques linéaires :

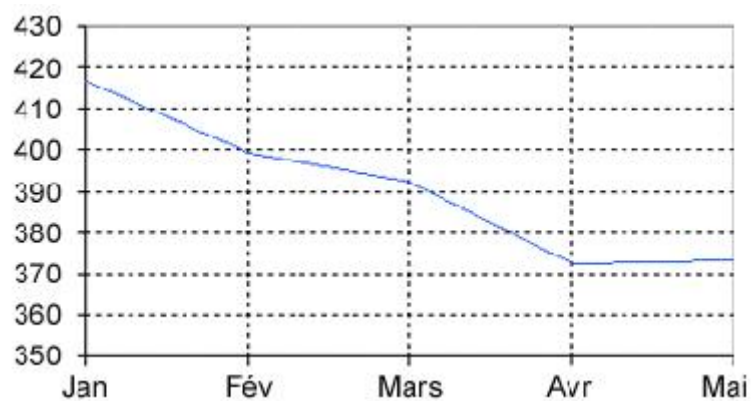


FIG. 2.3 – Graphique linéaire simple

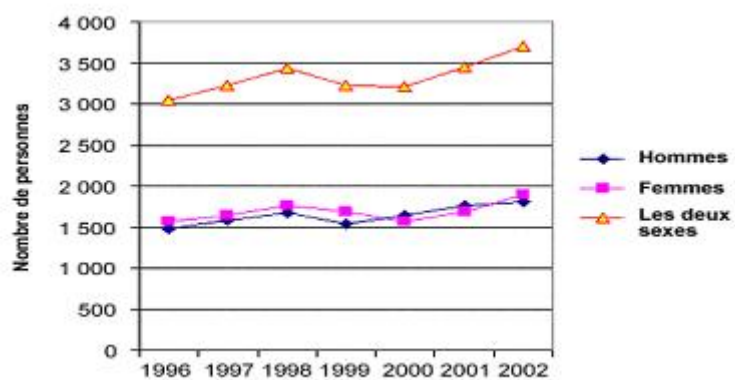


FIG. 2.4 – Graphique linéaire multiple

5. Nuage de points :

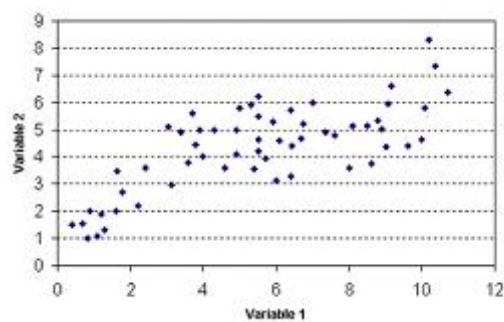


FIG. 2.5 – Nuage de points

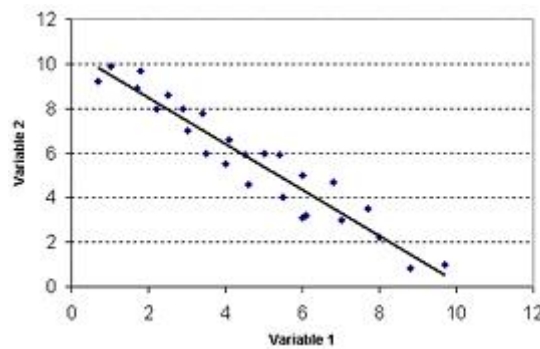


FIG. 2.6 – Nuages de points avec aide à la visualisation

Remarque : il peut exister des dépendances dans les nuages de points comme dans la figure précédente mais parfois le nuage peut faire apparaître des données très dispersées.

6. L'histogramme.

L'histogramme est un outil fréquemment utilisé, notamment pour résumer des données discrètes ou continues mesurées dans une échelle d'intervalle. Il est souvent employé pour montrer les caractéristiques principales de la distribution des données de façon pratique. Un histogramme sépare les valeurs possibles des données en classes ou groupes. Pour chaque groupe, un rectangle est construit dont la base correspond aux valeurs de ce groupe, et dont la taille du rectangle est proportionnelle au nombre d'observations dans le groupe. Cela signifie que les rectangles seront d'une hauteur différente. Un histogramme a une apparence semblable au graphique à barres verticales, mais lorsque les variables sont continues, il n'y a pas d'écart entre les barres. Lorsque les variables sont discrètes, on ne parle plus d'histogramme mais de diagramme à bâtons ou Bar Chart (figure 2.7). La figure 2.8 est quant à elle un bon exemple d'histogramme.

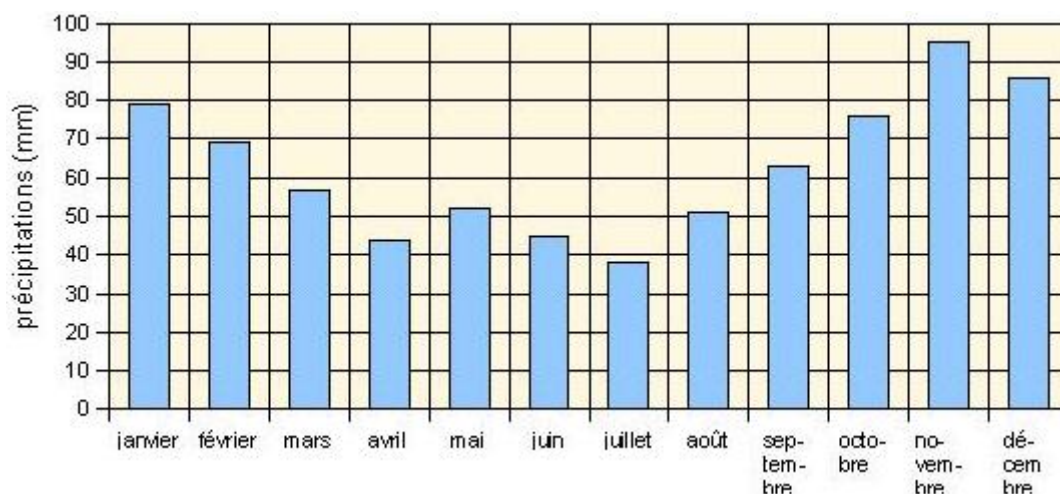


FIG. 2.7 – Diagramme à bâtons ou Bar Chart

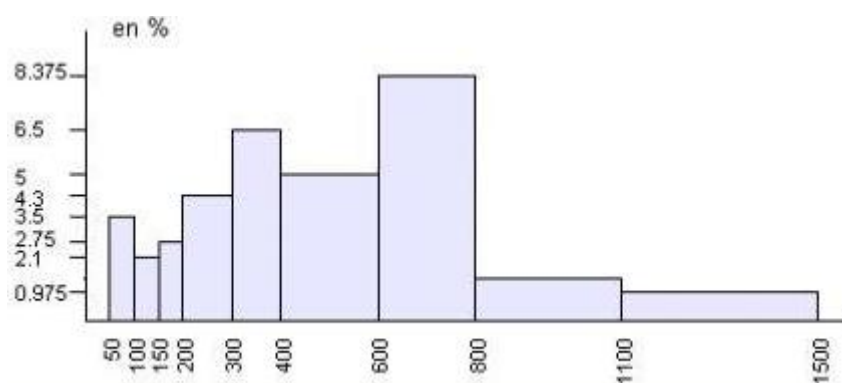


FIG. 2.8 – Histogramme

Les tableaux

Les tableaux, comme les graphiques regroupent des données essentiellement quantitatives. Il n'y a pas de types différents de tableaux mais seulement des présentations différentes suivant leur ordonnancement : Tableaux ordonnés (croissants, décroissants), tableaux désordonnés, tableaux semi-ordonnés.

2.2 Visualisations boursières pour téléphones mobiles sur Sites Wap

Nous allons à présent parcourir les différentes méthodes d'affichages d'information visuelle sur les téléphones mobiles dans le domaine de la bourse. Pour cela, nous examinerons quelques images prises sur des sites Wap financiers tels que <http://wap.boursorama.com>

, <http://www.firstinvest.waptoo.com/>,

<http://wireless.stockpoint.com>,... Les sites Wap sont les sites accessibles par les téléphones mobiles c'est à dire qu'ils sont à la dimension des petits écrans de GSM. Ces sites ne sont pas construit en HTML comme les sites Web mais en WML qui est un cousin de HTML. Il faut savoir aussi que certains sites Web boursier bien connus comme le site de NASDAQ (National Association of Securities Dealers-Automated Quotation : Réseau télématique d'information et de cotation sur le marché américain des valeurs mobilières des entreprises nouvellement créées et à fort potentiel de croissance) n'offre pas de service Wap. Les différentes méthodes de visualisation rencontrées sur les sites Wap dans le domaine de la bourse sont les tableaux et les graphiques.

2.2.1 Les tableaux

Les tableaux sont très utilisés pour représenter les valeurs des différents indices boursiers. Chaque ligne du tableau représente un indice et ses différentes valeurs telles que la valeur la plus basse, la valeur la plus haute, le pourcentage de variation,... Sur les sites Wap, les tableaux ne sont pas représentés de la même manière que sur des sites Web. Afficher un tableau d'indices (Dow Jones, CAC40,...) ne serait pas possible sur une carte ou bien serait illisible. La technique utilisée en général sur les sites Wap est d'afficher d'abord une carte avec une série de liens représentant les différents indices boursiers et quand l'utilisateur a choisi son indice, une autre carte s'affiche avec les différentes informations sur l'indice (ouverture, valeur la plus basse,...). Cela revient à choisir dans la première carte quelle ligne du tableau on veut consulter.

La plupart des sites Wap rencontrés utilisent une liste de liens pour représenter les tableaux. Quand un utilisateur choisit par exemple de voir les valeurs des indices du marché boursier (figure 2.9), une nouvelle carte s'affiche avec une liste de liens (figure 2.10). Chaque lien représente une ligne dans le tableau, il suffit juste à l'utilisateur de choisir l'indice (donc la ligne) qu'il veut consulter. Remarquons qu'un ascenseur est disponible sur le côté droit de l'écran pour permettre de se déplacer dans la liste des indices car celle-ci est trop longue pour apparaître entièrement à l'écran. Ensuite, dès que l'utilisateur a cliqué sur un lien, une autre carte apparaît à l'écran avec un graphique donnant la valeur de l'indice choisi préalablement (en abscisse le temps et en ordonnée la valeur)(figure 2.11). Pour arriver à consulter le reste des informations de la ligne, on trouve sur la deuxième carte un lien appelé détail qui va, après avoir cliqué dessus, donner les informations sur l'indice (figure 2.12).



FIG. 2.9 – Menu principal (www.wap.boursorama.com)



FIG. 2.10 – Carte 1, choix de l'indice. Correspond au choix d'une ligne dans le tableau.



FIG. 2.11 – Carte 2, lien détail vers le reste des informations de l'indice.



FIG. 2.12 – Carte 3, ligne du tableau.

Voyons un autre exemple d'un site Wap avec des tableaux (<http://www.cprbourse.tm.fr/wap>). Dans celui-ci, aucun graphique n'est fourni avec le tableau, c'est pourquoi quand l'utilisateur a choisi son indice (figure 2.13), les informations apparaissent tout de suite dans une nouvelle carte (figure 2.14). L'utilisation d'abréviation est très courante car c'est un gain de place qui ne peut s'avérer négligeable dans le cas des téléphones portables. Ouverture est dès lors remplacé par Ouv (figure 2.14).



FIG. 2.13 – Choix de l'indice.



FIG. 2.14 – Informations sur l'indice.

2.2.2 Les graphiques

La plupart des graphiques trouvés sur les sites Wap sont des graphiques avec des courbes simples. Certains sont plus lisibles que d'autres, certains utilisent des couleurs mais la plupart sont en noir et blanc, certains graphiques n'ont pas d'axes,... On peut aussi remarquer que les Bar Charts ne sont pas beaucoup utilisés sur les sites alors que ces représentations valent sans aucun doute les représentations en courbe.

La figure 2.15 représente la visualisation d'un graphe sur le site wap de boursorama. La courbe est presque une droite car les valeurs du cours ont été prises lors de la dernière heure et donc elles n'ont pas beaucoup varié. Par contre la figure 2.16 donne l'historique de la même action mais cela sur plusieurs heures permettant ainsi de voir les différentes variations de l'action. On remarque que l'axe des ordonnées est à droite du graphe sur ce site.



FIG. 2.15 – Cours de l'action SONY lors de la dernière heure.



FIG. 2.16 – Historique du cours de l'action SONY lors des dernières heures.

A l'inverse, sur le site de moneyextra, l'axe des ordonnées est quant à lui à gauche du graphique, représentation la plus classique d'un graphique.

FIG. 2.17 – Cour de l'action SIE sur le site de <http://wap.moneyextra.com>.FIG. 2.18 – Cour d'une action sur le site de www.firsinvest.waptoo.com

2.3 Application existante

2.3.1 Midcast

Développé par HillCast Technologies, Midcast est une application pour les réseaux Wireless dans le domaine des services financiers. Elle met les données de la bourse dans les mains des observateurs du marché en leur délivrant des informations boursières avec des graphes et des tableaux sur des téléphones mobiles ou des PDA.

MidCast est une application Java qui donne les changements de prix des actions, volumes, dernière quantité échangée, la valeur la plus haute et la plus basse d'une action, le jour d'ouverture,...

Pour accéder aux informations boursières, l'utilisateur sélectionne MidCast dans le menu des applications de son téléphone ou de son PDA. Le logiciel procède alors à une demande de connexion auprès du serveur MidCast. Une fois que la connexion est opé-

rationnelle et la valeur d'une action demandée, des mises à jour pour ces actions sont automatiquement envoyées à l'application.

Le logiciel est en contact permanent avec le serveur qui lui distribue les données financières incluant les valeurs, les détails des actions, des tableaux et des graphes pour les afficher sur le GSM ou bien le PDA.

MidCast possède une interface graphique avec des tableaux journaliers et des historiques. A présent il peut opérer sur 100 dispositifs dont les téléphones mobiles et les PDA.

Dans Midcast, l'utilisateur peut visualiser un tableau reprenant les différentes actions qu'il possède, c'est-à-dire son portefeuille (figure 2.19). Bien-sûr le plus important ici est de voir comment Midcast affiche ces informations. Le tableau possède cinq colonnes. La première colonne n'est pas quantitative, elle ne possède que des flèches soit vers le haut soit vers le bas qui représente le fait qu'une action ait variée positivement (vers le haut) ou bien négativement (vers le bas), la deuxième est le symbole de la société cotée, la troisième la valeur actuelle des actions, la quatrième le taux de variation et la cinquième le volume. On peut ainsi remarquer qu'une flèche vers le haut dans la première colonne est associée à un taux de variation positif. Les deux, le taux et la flèche, sont en vert. Inversement, à une flèche vers le bas est associé un taux de variation négatif et les deux sont en rouge.

T	ID	PRICE	CHG	VOL
↑	AOL	49.50	0.70	3186
↓	AMET	42.96	-0.70	2085
↑	DELL	26.00	0.70	7488
↓	MSFT	71.00	-0.11	8252
↑	MOT	22.25	0.00	1572

At the bottom of the screen are buttons labeled 'BACK' and 'ADD'.

FIG. 2.19 – Tableau comprenant les actions du portefeuille d'un utilisateur.

Les Candlesticks sont également présents dans cette application (figure 2.20). Un candlestick peut ainsi être affiché à l'écran pour chaque action. Rappelons que la fin d'une ligne vers le haut représente le cours le plus haut de l'action, la fin de la ligne vers le bas le cours le plus bas, la fin du rectangle vers le haut la clôture et la fin du rectangle vers le bas l'ouverture (tout ceci pour une certaine période). Aussi dans cette représentation le vert signifie que le cours a augmenté durant la période, le rouge que le cours a baissé.



FIG. 2.20 – Candlestick d'une action.

L'utilisateur peut aussi visualiser un graphe donnant le cour d'une action (en ordonnée) lors des trois dernières minutes avant le moment présent (figure 2.21).



FIG. 2.21 – Graphe d'une action lors des trois dernières minutes.

Finalement, on peut aussi afficher plus d'informations sur une action. Ceci pourrait être interprété comme la ligne d'un tableau général des actions. L'action devant être visualisée serait en fait l'entrée d'une ligne dans le tableau général (DELL dans la figure 2.22).



FIG. 2.22 – Informations boursières d'une action.

2.4 Conclusion

Comme on peut le voir, les téléphones portables sont très touchés par le domaine boursier. Le nombre de services boursiers offerts aux utilisateurs de GSM ne cesse d'augmenter. Ceci est dû sans aucun doute au nombre d'abonnés de téléphones mobiles qui ne cesse d'augmenter. La bourse étant de plus un domaine qui nécessite une constante attention, les GSM ont donc l'avantage de permettre à tout le monde de s'informer sur ce domaine à tout moment en temps réel et en tout lieu.

Cependant les graphes proposés sur les sites Wap ne sont guères lisibles et parfois presque inutiles. Les applications sont quant à elle plus intéressantes mais sont beaucoup moins complètes que pour les PDAs. De plus, il en existe moins que pour les PDAs, toutes doivent être téléchargées sur le Net et elles ne sont pas disponibles pour tous les GSM.

Chapitre 3

Analyse des besoins

Plus d'un milliard de personnes, près d'un sixième de la population mondiale, utilisent maintenant les téléphones mobiles. Cette étape historique a été atteinte 12 ans seulement après le lancement des premiers réseaux. Le GSM (Global System for Mobile communication), maintenant adopté dans plus de 200 pays et territoires, est devenu la norme globale unique des communications mobiles. Quatre-vingt pour cent de tous les nouveaux clients du réseau de téléphonie mobile ont choisi le GSM de sorte que le nombre de téléphones mobiles dépasse maintenant sous tous les aspects les lignes téléphoniques fixes. Pour la deuxième année consécutive, le secteur GSM a inscrit autant de nouveaux clients pendant l'année que la deuxième technologie mobile en importance (CDMA) comptait d'utilisateurs au total à la fin de l'année. Le GSM a atteint une étape marquante au premier trimestre de cette année. Comptant plus de 970 millions d'utilisateurs à la fin de décembre et ayant enregistré en moyenne une augmentation de 15 millions d'utilisateurs par mois en 2003, la GSM Association, organe représentatif des principaux exploitants GSM mondiaux, a enregistré son milliardième utilisateur GSM connecté au cours de ce premier trimestre. Le tableau 3.1 montre le nombre d'abonnés mobiles en 2002 et 2003 et l'évolution.

Année	Nombre d'abonnés(en Millions)	Evolution
2003	1.392,9	19.5%
2002	1.165,0	22,1%

TAB. 3.1 – Evolution du nombre d'abonnés mobiles dans le monde

Dès lors, les GSM deviennent de plus en plus perfectionnés, proposant de plus en plus de fonctionnalités telles que appareil photo, caméra,... Aussi vu que les performances technologiques des GSM s'améliorent de jour en jour, il est possible à présent de faire des applications qui avant n'étaient même pas imaginables. Dans notre cas présent, l'amélioration technologique majeure est sans aucun doute que les écrans de la majorité des

téléphones actuels sont en couleurs, permettant ainsi de fournir à l'utilisateur plus d'informations sur un plus petit écran. Cependant malgré l'accroissement technologique, les téléphones mobiles ne pourront jamais afficher beaucoup d'informations sur leur écran car la taille d'un GSM restera toujours fort petite. Ainsi il faut imaginer de nouvelles techniques pour permettre à l'utilisateur de visualiser sur son téléphone mobile une information volumineuse, en particulier.

La suite va définir le type d'utilisateurs, l'environnement, ainsi que les différentes tâches de l'application suivant une liste de contrôle basée sur [Noirhomme].

3.1 Utilisateurs

3.1.1 Types d'utilisateur

Deux types d'utilisateurs vont apparaître pour le produit. Tout d'abord, les personnes possédant un ensemble d'actions en bourse ou même aucune action et voulant s'informer sur l'état du marché en général ou bien de certaines actions en particulier. Ensuite le deuxième type regroupe les personnes qui sont journalièrement confrontées au domaine boursier. Que ce soit des personnes jouant tout le temps en bourse ou bien des spécialistes comme des courtiers en bourse.

Le premier type d'utilisateur sera appelé "Utilisateur novice" et le deuxième "Utilisateur expert"

3.1.2 Les caractéristiques de l'utilisateur

Utilisateur novice

1. Attributs physique :

- Intervalle d'âge : aucun âge n'est requis pour cette application mais les personnes en dessous de 20 ans s'intéressent en général moins au domaine boursier que leurs aînés.
- Sexe : le sexe n'a pas d'importance dans cette application, le pourcentage d'hommes et de femmes est donc le même. Cependant, on remarque qu'il y a plus d'hommes que de femmes dans le domaine de la bourse.
- Incapacités physiques et handicaps : ce type d'utilisateurs ne doit pas souffrir d'un problème de vue important (surtout de près).

2. Attributs mentaux :

- Capacités particulières : aucune capacité intellectuelle n'est requise pour ces personnes.

- Handicap mental : ces personnes ne souffrent d'aucun handicap mental.
 - Motivation : la motivation est positive vis-à-vis de la tâche car celle-ci permet aux utilisateurs de s'informer sur leurs actions en bourse et sur la santé du marché en général.
3. Qualifications et connaissances :
- Expérience avec l'ordinateur et le GSM : les compétences de ces personnes pour les ordinateurs peuvent aller du simple débutant jusqu'à un utilisateur plus expérimenté avec les ordinateurs car en général toute personne a des connaissances de base dans le domaine de la bourse et le GSM est devenu une chose courante de nos jours. L'utilisateur n'a donc pas besoin d'être un expert en informatique pour utiliser l'application.
 - Qualification générale : toutes les qualifications peuvent aller pour ce type d'utilisateur, cependant ils ont des connaissances limitées dans le domaine boursier et ne connaissent que les principes de base.
 - Connaissance des langues : le Français est la langue maternelle de l'utilisateur mais celui-ci peut avoir des connaissances en Anglais.
4. Caractéristiques du travail :
- Fonction : consulter un ensemble d'actions.
 - Flexibilité dans le travail : malgré que la bourse ouvre à 9h et ferme à 16h, les utilisateurs peuvent décider du moment où ils utiliseront le produit car le service est ouvert en permanence.

Utilisateur expert

1. Attributs physique :
- Intervalle d'âge : l'intervalle n'est pas limité comme dans le cas du premier type d'utilisateur mais les personnes en dessous de 20 ans ne sont en général pas dans cette catégorie.
 - Sexe : le sexe n'a pas d'importance dans cette application, le pourcentage d'hommes et de femmes est donc le même. Cependant, on remarque qu'il y a plus d'hommes que de femmes dans le domaine de la bourse.
 - Incapacités physiques et handicaps : ce type d'utilisateurs ne doit pas souffrir d'un problème de vue important (surtout de près).
2. Attributs mentaux :
- Capacités particulières : les experts ont plus de connaissances sur le domaine boursier et donc sont déjà pour la plupart familiarisés avec les visualisations d'informations classiques.
 - Handicap mental : ces personnes ne souffrent d'aucun handicap mental.

- Motivation : la motivation est positive pour la tâche car celle-ci leurs est utile dans leur travail quotidien.
- 3. Qualifications et connaissances :
 - Formation et expérience dans le processus et les méthodes qu'utilisent les tâches :
 - Expérience avec l'ordinateur et le GSM : ce type d'utilisateurs a de bonnes connaissances en informatique et dans le domaine de la bourse car il va régulièrement sur les sites boursiers pour son métier. Sa connaissance pour les GSM est normale comme tout le monde..
 - Qualification générale : qualification avancée dans le domaine de la bourse car ils exercent un métier en rapport avec la bourse.
 - Connaissance des langues : le Français est la première langue de l'utilisateur mais celui-ci parlera aussi l'Anglais un peu ou couramment.
- 4. Caractéristiques du travail :
 - Fonction : consulter un ensemble d'actions.
 - Flexibilité dans le travail : malgré que la bourse ouvre à 9h et ferme à 16h, les personnes très familières avec la bourse peuvent décider du moment où ils utiliseront le produit car le service est ouvert en permanence.

3.1.3 Environnement

Les deux types d'utilisateurs vont être regroupés ici car l'environnement est le même pour ces deux types que ce soit des utilisateurs novices ou experts.

Localisation du produit

Le produit peut être utilisé partout du moment qu'une connexion Wap est disponible à l'endroit où on veut l'utiliser.

Position

Le produit peut être utilisé dans n'importe quelle position, que ce soit assis, debout ou même en marchant.

Matériel

- Hardware de base : le produit sera utilisé exclusivement sur des téléphones portables couleurs.
- Interface : l'écran de GSM sera l'interface et l'utilisateur interagira avec le GSM grâce aux touches.
- Connexion au réseau : la connexion au réseau sera les connexion WAP.

Logiciel et système d'exploitation

- Software requis pour exécuter le produit : tous les systèmes d'exploitation sont permis du moment qu'ils acceptent les applications Java.
- Software portable : aucun autre software n'est requis pour ce produit mais l'application pourra être complétée par d'autres software pour acheter et vendre des actions car le produit ne le permet pas, il permet juste de pouvoir consulter des actions.

3.1.4 Tâches

Liste des tâches

Remarque : les tâches sont les mêmes pour les deux types d'utilisateurs.

- Consulter les cours moyens d'un ensemble d'actions sur un intervalle de 1 mois ou 3 mois.
- Consulter les variations journalières moyennes des cours d'un ensemble d'actions sur un intervalle de 1 mois ou 3 mois.

Caractéristiques de la tâche

Remarque : les deux tâches ont comme différence uniquement leur but (l'une est sur les cours des actions, l'autre sur les variations et donc elles vont être présentées en même temps).

- But de la tâche : pour la première tâche, le but est de consulter les cours moyens d'un ensemble d'actions sur 1 mois ou 3 mois. Pour la deuxième, le but est de consulter les variations moyennes des cours d'un ensemble d'actions sur 1 mois ou 3 mois.
- Alternative : les utilisateurs peuvent faire ces tâches en consultant des sites Web boursiers, des sites Wap boursiers ou bien des applications existantes.
- Sorties : sur l'écran du GSM apparaît les résultats sur une des visualisations proposées et choisies par l'utilisateur.
- Fréquence de la tâche : une fois par jour ou par semaine pour l'utilisateur novice. Plusieurs fois par jour pour l'utilisateur expert.
- Durée de la tâche : pour les deux tâches, la durée est de l'ordre de quelques minutes (en moyenne 5 minutes).
- Flexibilité de la tâche : il n'y a aucune séquence prédéfinie, l'utilisateur peut choisir n'importe quelle visualisation pour voir les résultats. Cependant il devra respecter un ordre logique, à savoir d'abord choisir l'intervalle de temps, les cours ou les variations des actions et le nombre de clusters avant de visualiser les résultats.
- Compétences et connaissances : pas de compétences particulières.
- Dépendances de la tâche : une connexion WAP au réseau est requise pour pouvoir récupérer les différents résultats.

- Tâches jointes : il n'existe aucune tâche jointe.

Chapitre 4

Limites des téléphones mobiles

Au contraire des ordinateurs de bureau ou des ordinateurs portables, les téléphones mobiles possèdent beaucoup de contraintes pour la programmation telles que la taille de l'écran, la mémoire disponible, le nombre de couleurs, la vitesse du processeur... Voyons de plus près chacune de ces contraintes.

4.1 Limitation par rapport à l'écran

La taille de l'écran est une contrainte très importante pour la visualisation de données. Plus l'écran est grand, plus il est facile de visualiser une plus grande quantité d'informations à la fois. Le tableau 4.1 reprend les tailles des écrans de différents Nokia.

Nokia	taille de l'écran
3100	128 x 128 pixels
3510i	96 x 65 pixels
3660	176 x 208 pixels
6100	128 x 128 pixels
6600	176 x 208 pixels
6800	128 x 128 pixels
7610	176 x 208 pixels
9500	640 x 200 pixels
N-Gage	176 x 208 pixels

TAB. 4.1 – Tailles des écrans de certains Nokia.

Néanmoins nous pouvons constater que la taille des écrans de GSM devient de plus en plus grande et que ces écrans deviennent de plus en plus nettes. Mais le but d'un téléphone portable étant d'être le plus petit possible pour pouvoir être porté sans difficulté par son propriétaire, la taille des écrans restera donc toujours un problème. Il faut donc profiter

au maximum de tout l'écran et d'utiliser les couleurs pour représenter des informations qu'on ne peut afficher en plus sur l'écran. Bien entendu on ne pourra jamais atteindre la même précision pour les graphiques que pour les PDA par exemple.

4.2 Limitation par rapport à la mémoire

Le manque de mémoire est souvent aussi un problème pour programmer facilement. Surtout pour cette application qui utilise un grand nombre de données. L'incapacité à créer une base de données comprenant toutes les informations nécessaires rend la programmation beaucoup plus difficile. Il existe néanmoins sur la plupart des nouveaux téléphones mobiles la possibilité d'enregistrer des données mais la capacité mémorielle reste très réduite. Le tableau 4.2 montre la capacité mémoire de certains téléphones Nokia.

Nokia	Mémoire
3100	450 Ko
3510i	180 Ko
3660	4 Mo + carte mémoire
6100	725 Ko
6600	6 Mo+carte mémoire
6800	5,2 Mo
7610	40 Mo + carte mémoire
9500	90 Mo + carte mémoire
N-Gage	4 Mo + carte mémoire

TAB. 4.2 – Capacité mémoire de certains Nokia.

4.3 Limitation par rapport à la vitesse du processeur

La vitesse des processeurs des téléphones mobiles étant beaucoup plus lente que celle des ordinateurs traditionnels, il est beaucoup plus difficile de faire tourner des applications demandant de longs calculs. De plus, le temps de réponse d'un programme est une des caractéristiques les plus importantes pour celui-ci. Il convient donc de faire exécuter par le téléphone mobile le moins de code possible. Si il faut par exemple faire afficher par le téléphone un dessin complexe, le mieux est de faire ce dessin sur un ordinateur normal au préalable pour ensuite le faire afficher par le téléphone au lieu de laisser celui-ci le dessiner petit à petit.

4.4 Problème de navigation

Les téléphones portables n'ont pas de facilité de navigation comme les ordinateurs qui possèdent des souris ou bien les PDAs et leur stylo. La navigation sur GSM se fait par le clavier numérique et deux trois boutons supplémentaires. On peut dès lors s'apercevoir que sur un PC, naviguer rien qu'avec le clavier ne facilite en rien son utilisation. Malgré tout, les particuliers sont de plus en plus familiers avec la navigation des téléphones portables vu que l'utilisation de ceux-ci est devenue chose courante. La programmation d'une application doit quant à elle faire attention à avoir une navigation fluide pour ne pas que l'utilisateur se perde. Les boutons doivent clairement faire ce qui est indiqué qu'ils font et l'application doit toujours permettre à l'utilisateur de revenir en arrière. De plus les différents boutons varient de GSM en GSM. Cette contrainte est quand même moins difficile à surmonter mais elle reste d'une importance capitale pour la cohérence de l'application.

4.5 Conclusion

Les téléphones portables ont donc beaucoup plus de contraintes que sur les autres "hardwares" tels que les ordinateurs ou les PDAs. Pour programmer sur GSM, il faut donc connaître au mieux toutes ces contraintes pour pouvoir les contourner. Celles-ci portent d'ailleurs sur des caractéristiques essentielles de tout "hardware", que ce soit la mémoire, la vitesse du processeur ou bien la taille de l'écran. D'autres contraintes sont d'un autre niveau et portent sur la facilité d'utilisation, tout spécialement la navigation dans l'application.

La technologie peut par contre modifier la donne et faciliter la vie des programmeurs car les GSM deviennent de plus en plus perfectionnés repoussant sans cesse les limites de ceux-ci. Certaines caractéristiques des GSM sont améliorées mais d'autres ne pourront jamais atteindre celle des PDAs ou des ordinateurs, ceci à cause d'une des fonctionnalités principales des GSM, leur portabilité et donc leur taille réduite.

Chapitre 5

Classification

La classification est une méthode d'analyse de données qui vise à regrouper en classes les données les plus semblables. Elle est aussi appelée classification non supervisée ou clustering, pour la distinguer d'une classification dans des classes connues à priori. Le principe général est de maximiser la distance entre classes différentes et de minimiser la distance interclasse.

Les principales applications de la classification sont le traitement d'images, l'identification de modèles (pattern recognition) les Sciences Economiques (aide à la recherche pour le marché), la classification de documents sur le Web,... Voici des exemples d'application de la classification :

- Marketing : aider les vendeurs à découvrir des groupes distincts dans leurs bases de données des clients, et puis employer ces connaissances pour développer des programmes de vente pour des marchés bien spécifiques.
- Cultivation de la terre : identification des domaines de la cultivation semblable de la terre dans une base de données d'observation de la terre
- City-Planning : identifier des groupes de maisons selon leur type, leur valeur et leur localité.
- Études des tremblements de terre : Les épicentres des tremblements de terre observés devraient être classifiés selon les continents.

Il existe plusieurs méthodes de clustering [Everitt]. Les méthodes hiérarchiques agglomèrent les données à partir des données individuelles en commençant par les grouper deux à deux. Ceci a le désavantage d'être très lourd au point de vue temps de calcul lorsque les données sont un peu nombreuses. Par contre, les méthodes de nuées dynamiques partent d'une partition à priori et l'améliorent par optimisations successives. Ce processus converge, en général, rapidement. L'inconvénient est qu'il faut fixer le nombre de

classes dès le départ. Cependant il est possible d'utiliser en complément une méthode qui optimise le nombre de classes [2]. Vu sa rapidité, nous avons choisi, ici, une telle méthode. Plus précisément, nous avons utilisé la méthode du K-Means ou des centres mobiles.

5.1 La méthode des K-Means

Son principe est assez simple. K est donné au départ et représente le nombre de classes. A l'initialisation, les K premières données vont être les K premiers centres mobiles (il est possible aussi de choisir ces premiers centres en fonction d'une connaissance a priori). Chaque point autre que les centres va ensuite être assigné au centre le plus proche, c'est-à-dire dont la distance est moindre. L'adjonction de points va changer les centres effectifs si bien que les centres des classes ainsi constituées sont recalculés. Chaque point est ensuite réassigner au centre le plus proche et ainsi de suite. Le processus s'arrête si à l'itération $i + 1$, pour chaque classe, centre et points de la classe sont les mêmes qu'à l'itération i . La figure 5.1 donne un exemple du principe de la méthode des K-Means.

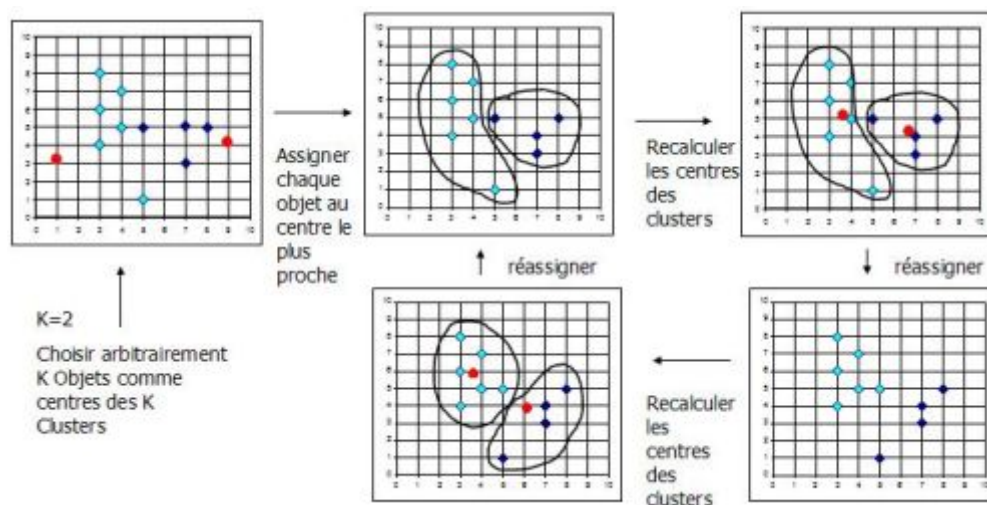


FIG. 5.1 – Exemple du principe des K-Means où $K=2$

Pour les actions boursières, nous avons considéré comme donnée la valeur moyenne de l'action pour une certaine période. Nous avons aussi considéré la variation de la valeur pour cette période. Nous avons ensuite appliqué la méthode du K-means pour regrouper les actions en classes, en utilisant la distance euclidienne. Ceci nous permet donc de réduire l'information à quelques classes. Par exemple, nous avons utilisé la méthode pour regrouper 40 actions, à partir de leur taux de variation entre deux jours ouvrables. Une fois les actions classifiées, l'utilisateur peut voir quels groupes d'actions diminuent fortement ou quels groupes sont en hausse progressive.

5.2 Exemple de calcul des K-Means

Donnée : 2,4,10,12,3,20,30,11,25, $k=2$

Assigné les centres de façon arbitraire : $m1 = 3, m2 = 4$

$K1=2,3, K2=4,10,12,20,30,11,25, m1 = 2.5, m2 = 16$

$K1=2,3,4, K2=10,12,20,30,11,25, m1 = 3, m2 = 18$

$K1=2,3,4,10, K2=12,20,30,11,25, m1 = 4.75, m2 = 19.6$

$K1=2,3,4,10,11,12, K2=20,30,25, m1 = 7, m2 = 25$

Stop car aucune assignation supplémentaire ne changera les centres.

5.3 Apport de cette méthode

Pour les techniques de visualisations de l'information, la classification permet de réarranger les attributs des données entre elles pour aider un grand nombre de ces techniques de visualisations ; des techniques comme les coordonnées parallèles,... L'idée de base est alors de regrouper les attributs des données qui ont un comportement similaires les uns à coté des autres pour offrir une présentation synthétique. La similarité des variables aléatoires est un prérequis pour trouver l'arrangement optimal à une ou deux dimensions des attributs des données[Ankerst]. Le terme "attribut" est le terme utilisé dans le domaine des bases de données.

Ensuite, la classification permet, comme déjà dit, de regrouper un ensemble de données et donc de diminuer la quantité d'information à présenter à l'utilisateur. Sur un même écran de GSM, on peut dès lors afficher des informations sur un ensemble d'actions sous forme de clusters. L'utilisateur peut ainsi se connecter à Internet et télécharger les détails de ses actions sous forme de clusters. Il peut ensuite se déconnecter d'Internet et afficher à l'écran de son GSM les différents clusters et leur valeur. Les clusters seront donc composés d'un ensemble d'action qui ont un comportement semblable c'est-à-dire des valeurs ou des variations semblables. Cependant les informations données avec le cluster ne seront qu'une moyenne des actions contenues dans un cluster. Par exemple, la valeur moyenne de 5 actions sur un mois représente la valeur du premier cluster et ainsi de suite. Mais pour les utilisateurs ne voulant pas passer beaucoup de temps sur Internet et ne voulant qu'une idée globale de l'évolution de l'ensemble de leur portefeuille d'actions, cette technique est très intéressante.

5.4 Conclusion

La classification non-supervisée est donc une méthode permettant de diminuer le nombre de données et de pouvoir donner des informations sur un grand ensemble de données en les regroupant en classes suivant certains critères. Il existe plusieurs méthodes de classification mais celle utilisée dans l'application est la technique des K-Means ou des centres mobiles. Un des désavantages de cette application est que l'on doit connaître le nombre de classes à priori. Cependant il existe des algorithmes d'optimisations qui permettent de calculer le nombre de classes optimal [Hardy]. La méthode des K-Means peut donc être affinée et inclure un calcul permettant de connaître le nombre optimal de classes.

L'utilisation de cette technique pour les GSM permettrait donc d'afficher sur un petit écran des informations relatives à un grand ensemble de données. Les informations sont des valeurs représentant chaque classe et en conséquence elles ne sont que des moyennes des données composant leur classe respective.

Chapitre 6

Les visualisations boursières de l'application

6.1 Introduction

La partie précédente se concentrant plus sur le traitement des données, cette section va se concentrer sur l'affichage des données. Ayant à présent une série de données réparties en classes, nous allons étudier les moyens pour afficher de telles données. Quatre visualisations boursières vont être présentées, deux visualisations classiques, élémentaires et deux visualisations plus récentes.

6.2 Visualisations Classiques

Les deux visualisations de base sur lesquelles nous avons portées notre attention sont le Bar Chart et le Tableau. Celles-ci sont bien entendu des visualisations courantes pour les personnes s'intéressant à la bourse.

6.2.1 Le Bar Chart

Le Bar Chart est une visualisation particulièrement intéressante dans le domaine de la bourse car en un coup d'oeil, l'utilisateur peut voir le maximum et le minimum d'un ensemble de données. Un Bar Chart est par définition un type de graphique dans lequel les différentes valeurs sont représentées par des barres verticales. Ici, chaque rectangle représente la valeur assignée à une classe. Souvent, cependant, il est nécessaire de représenter des nombres négatifs. C'est le cas lorsqu'on affiche des variations et non des valeurs moyennes. Comment alors faire pour les représenter en tenant compte de l'espace disponible et en veillant à respecter le soucis de lisibilité ? Deux types de Bar Charts vont être présentés pour répondre à cette question.

Bar Chart de type 1

Le premier Bar Chart affiche les nombres positifs avec des barres partant du bas de l'écran vers le haut mais affiche les nombres négatifs avec des barres partant du haut et allant vers le bas. Cela permet à l'utilisateur de voir rapidement combien de négatifs sont présents sur le graphe et aussi de gagner de la place sur l'écran. La barre maximum est coloré en vert et la barre minimum en rouge.

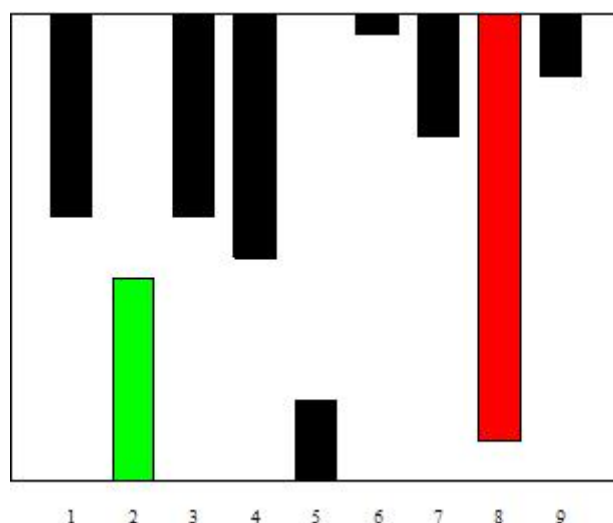


FIG. 6.1 – Bar Chart de type 1

Sur téléphone mobile, cela donne :

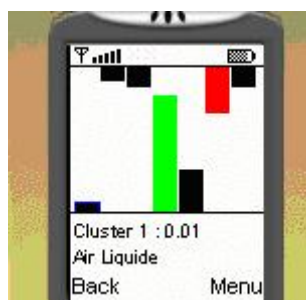


FIG. 6.2 – Bar Chart de type 1 sur téléphone mobile

Bar Chart de type 2

Le deuxième quant à lui, affiche des barres rectangulaires du bas vers le haut que ce soit pour les positifs ou bien pour les négatifs. Cependant un signe '+' est dessiné au dessus d'une barre si elle représente un nombre positif et un signe '-' au dessus d'une barre pour

un nombre négatif. Ici également le but est de permettre à l'utilisateur de voir les négatifs plus rapidement et de gagner de l'espace sur l'écran. Tout comme le Bar Chart de type 1, le maximum est en vert et le minimum en rouge.

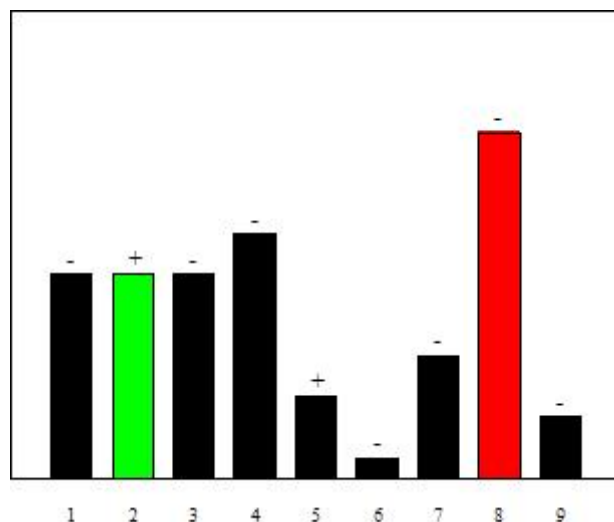


FIG. 6.3 – Bar Chart de type 2

Sur téléphone mobile, cela donne :

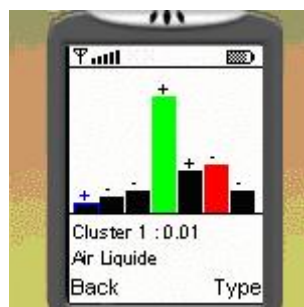


FIG. 6.4 – Bar Chart de type 2 sur téléphone mobile

Contenu de chaque cluster

L'utilisateur doit savoir malgré tout quelles actions sont dans quels clusters sinon l'application n'aurait pas beaucoup de sens. Mais comment afficher à l'utilisateur les contenus de chaque cluster sur l'écran du GSM en même temps que le Bar Chart. L'idée est d'utiliser les flèches directionnelles des téléphones portables pour permettre à l'utilisateur de faire défiler l'écran (flèche haut et bas) et de voir apparaître en bas le contenu de chaque cluster. Les flèches gauche et droite permettent de changer de clusters. Le contour de la

barre sur laquelle se trouve l'utilisateur clignote en bleu pour permettre à l'utilisateur de voir sur quel cluster il est.



FIG. 6.5 – Bar Chart de type 1 après avoir fait défiler l'écran vers le bas pour consulter le contenu du cluster 1

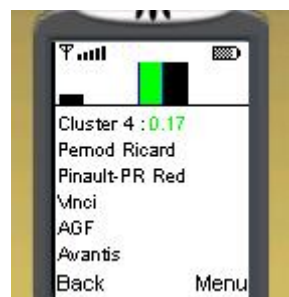


FIG. 6.6 – Bar Chart de type 1 avec contenu du cluster 4

6.2.2 Le tableau

Le tableau est peut-être la visualisation boursière la plus connue en dehors de la courbe traditionnelle. Le tableau permet surtout à l'utilisateur d'avoir rapidement accès à des données précises et a l'avantage de ne pas demander beaucoup de connaissances pour pouvoir le lire. Les numéros de la première colonne sont les numéros des classes, la deuxième colonne représente la variation correspondante.

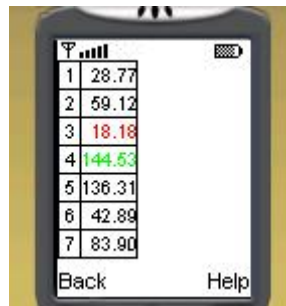


FIG. 6.7 – Représentation d'un tableau dans l'application

6.3 Les nouvelles visualisations

Nous allons présenter deux visualisations moins classiques. Nous avons adapté aux téléphones mobiles des techniques proposées par la communauté de la visualisation d'information ([Weber] et [Keim]). Ces visualisations permettent de représenter l'évolution de la valeur sur une période donnée et plus seulement une valeur moyenne. Elles visent toutes les deux à exploiter au maximum la taille de l'écran.

6.3.1 Le pixel Bar Chart

Le Pixel Bar Chart est une généralisation du Bar Chart traditionnel. L'idée est de représenter chaque donnée par un simple pixel sur le Bar Chart. En coloriant ce pixel, on peut dès lors donner plus d'informations sur une donnée que dans le simple Bar Chart.

Cependant, comment arranger les pixels dans chaque barre ? L'idée est de séparer les données dans les barres grâce à un attribut et ensuite d'utiliser un ou deux attributs supplémentaires pour imposer un ordre dans chaque barre. La figure 6.8 traduit cette idée.

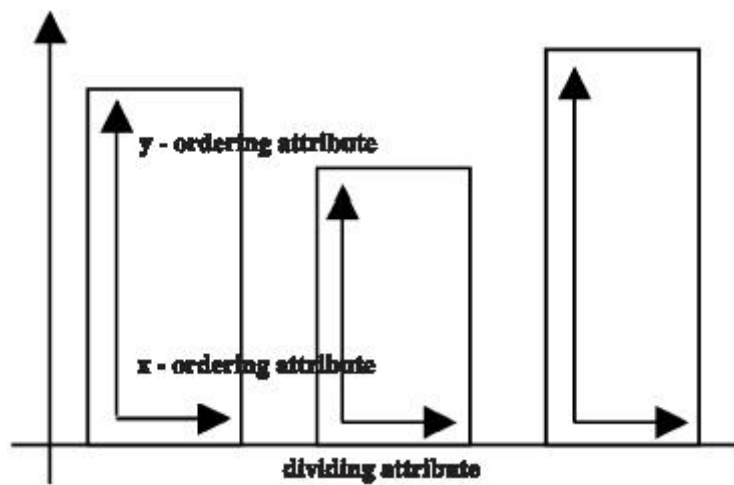


FIG. 6.8 – Principe du Pixel Bar Chart [Kei]

Un autre problème du Bar Chart est qu'une large portion de l'écran ne peut être utilisée à cause des différentes hauteurs de chaque barre. Mais avec un large ensemble de données nous voudrions avoir plus d'espace d'affichage disponible pour visualiser ces données. Une idée pour résoudre ce problème est d'utiliser des Bar Charts avec des barres de même hauteur à la place de Bar Charts avec des barres de même largeur. Les figures 6.9 et 6.10 représentent respectivement un Bar Chart avec des barres de même largeur et un Bar Chart avec des barres de même hauteur. Dans le premier cas, la valeur de la première barre est la hauteur de cette barre, dans le deuxième cas, la valeur est représentée par la largeur de cette barre.

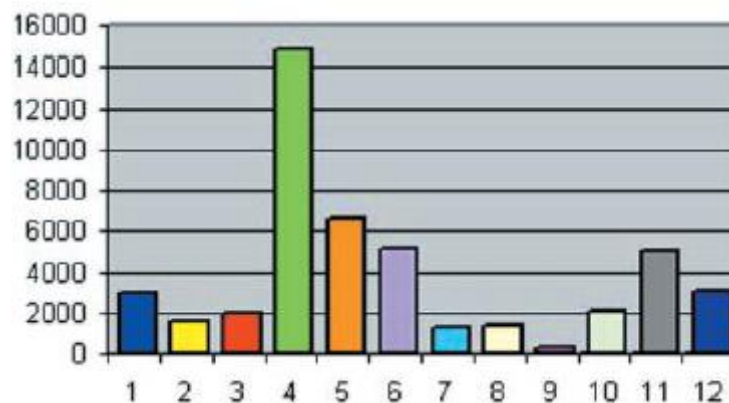


FIG. 6.9 – Bar Chart avec des barres de même largeur. Ce Bar Chart représente le nombre de clients pour 12 type de produits différents[Keim]

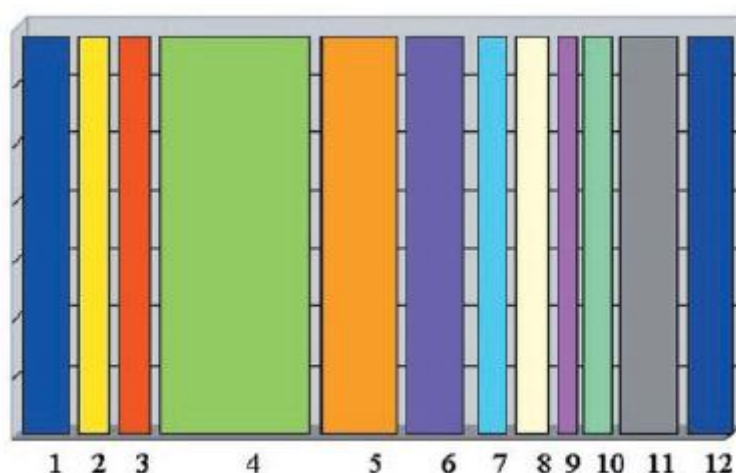


FIG. 6.10 – Bar Chart avec des barres de même hauteur. La largeur d'un rectangle représente ici le nombre de client pour le type de produit correspondant[Keim]

Maintenant appliquons notre idée du Pixel Bar Chart au Bar Chart des figures. Nous obtenons donc à la figure 6.11 un pixel Bar Chart avec des barres de même largeur tandis qu'à la figure 6.12, nous avons un Pixel Bar Chart avec des barres de même hauteur. On peut voir que cette dernière visualisation occupe tout l'écran fournissant aussi plus d'informations.

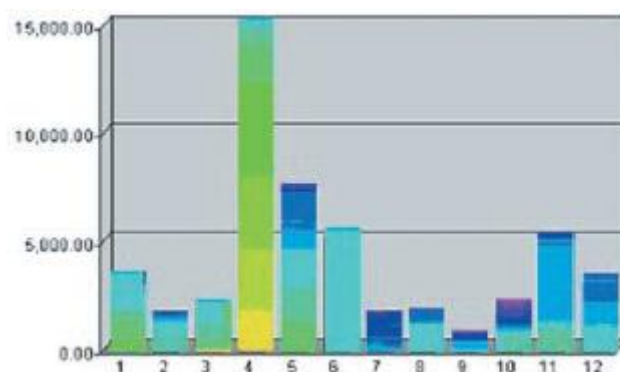


FIG. 6.11 – Pixel Bar Chart avec des barres de même largeur. L'attribut séparant est toujours le nombre de clients par type de produit. Le nombre de visite et la quantité d'argent donnée pour le produit sont les attributs x et y[Keim]

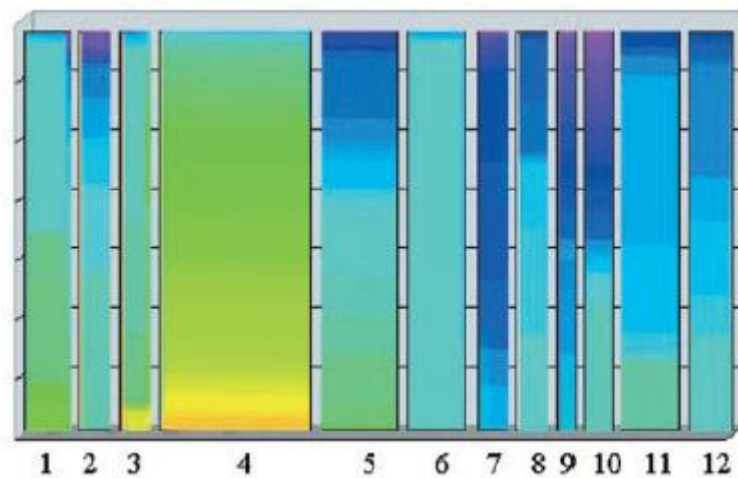


FIG. 6.12 – Pixel Bar Chart avec des barres de même largeur. Tout comme la figure précédente, l'attribut séparant est le nombre de clients par type et les attributs x et y sont le nombre de visite et la quantité d'argent donnée pour le produit[Keim]

Pour le téléphone mobile cependant, cette visualisation a été modifiée pour plus de facilité d'implémentation. Voyons de suite ce que cela donne sur téléphone mobile.

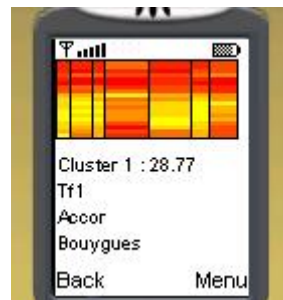


FIG. 6.13 – Pixel Bar Chart sur téléphone mobile

La largeur d'un rectangle représente la valeur du cluster correspondant c'est à dire que l'attribut séparant dans ce Pixel Bar Chart est la valeur de cluster. Les couleurs dans un rectangle (du bas vers le haut) représentent l'évolution du cluster au cours de l'intervalle choisi (ici, 1 mois). Nous avons choisi une couleur claire pour une valeur au-dessus de la moyenne du cluster pour l'intervalle et une couleur foncée pour une valeur basse. L'attribut x est donc le temps (il n'y pas d'attribut y). Les couleurs en bas des rectangles représentent donc les valeurs des clusters il y a un mois. La couleur permet de voir si la valeur à ce moment là est au-dessus (couleur clair, jaune), en-dessous de la moyenne (couleur foncée, rouge) ou près de la moyenne (couleur ni clair ni foncée, orange).

Tout comme pour le Bar Chart, le Pixel Bar Chart permet aussi à l'utilisateur de visualiser le contenu de chaque cluster (voir figure 6.14 et 6.15). Les flèches directionnelles ont les mêmes fonctions que pour le Bar Chart. Cependant, par manque de place, les rectangles ne possèdent pas de contour en bleu clignotant pour permettre à l'utilisateur de voir sur quel cluster il se trouve. Mais il est indiqué en dessous du Pixel Bar Chart le nom du cluster sur lequel l'utilisateur est.

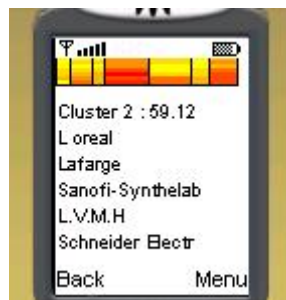


FIG. 6.14 – Pixel Bar Chart sur téléphone mobile avec contenu du cluster 2



FIG. 6.15 – Pixel Bar Chart sur téléphone mobile avec contenu du cluster 7

6.3.2 La spirale [Weber]

La spirale ne représente pas comme les autres visualisations présentées précédemment un ensemble de données, mais l'évolution de deux données et l'intérêt de cette visualisation est que l'on pourra comparer ces deux évolutions. Le principe de cette visualisation est de dessiner deux spirales, une intérieure et l'autre extérieure. Aucune échelle de temps n'est dessinée sur la spirale par manque de place et l'utilisateur ne pourra pas comparer précisément les deux clusters à une certaine date. Le but de cette visualisation est de comparer deux clusters sur une période de temps donnée mais de manière générale. On peut voir l'idée de cette visualisation sur la figure 6.16.

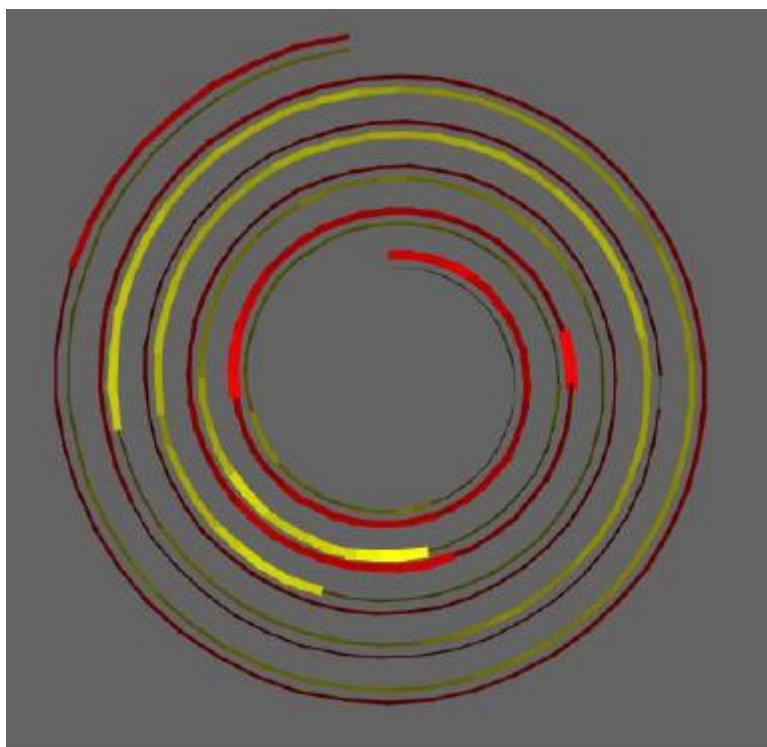


FIG. 6.16 – Spirale

Chaque spirale va être colorée et représentera l'évolution d'une donnée au cours du temps. Ainsi on pourra comparer les deux évolutions des deux données au cours du temps en comparant les couleurs. Le choix des couleurs est identique au choix fait pour le Pixel Bar Chart. Bien entendu, comme déjà dit, vu le peu de place disponible sur l'écran d'un GSM et une palette de couleurs très restreinte, cette évolution n'est qu'approximative et donne une idée globale des évolutions des deux données. La figure 6.17 montre cette visualisation sur téléphone mobile.



FIG. 6.17 – Visualisation de la spirale sur téléphone mobile

6.4 Conclusion

Nous avons vu dans ce chapitre quatre visualisations. Deux d'entre elles sont plus facile à utiliser et à comprendre mais présentent peu d'informations à l'utilisateur. Les deux autres par contre offrent plus de possibilités à l'utilisateur mais elles nécessitent plus de connaissances de la part des utilisateurs pour les comprendre. Ces quatre visualisations combinées à la technique du clustering sont les principales bases de notre application.

Chapitre 7

Implémentation

7.1 Choix du langage

7.1.1 Introduction

Dans le domaine de la programmation des mobiles, deux langages d'implémentation émergent principalement : java et C++. Bien qu'ils soient similaires du point de vue du code, leur distribution et utilisation sont assez différentes. Le langage utilisé dans l'application est Java pour une raison simple : ce langage nous était plus familier.

7.1.2 Java Micro Edition

La problématique multi-plate-formes et multi-périphériques de Java

Java 2 Micro Edition est une architecture technique dont le but est de fournir un socle de développement aux applications embarquées, l'intérêt étant de proposer tout la puissance d'un langage comme Java associé aux services proposés par une version bridée du Framework J2SE : J2ME. Les terminaux n'ayant pas les mêmes capacités en terme de ressources que les ordinateurs de bureau classiques (mémoire, disque et vitesse de calcul), la solution passe par la fourniture d'un environnement allégé afin de s'adapter aux différentes contraintes d'exécution. Cependant, comment faire en sorte d'intégrer la diversité de l'offre à un socle technique dont la cible n'est pas définie à priori ? La solution proposée par J2ME consiste à regrouper par catégories certaines familles de produits tout en proposant la possibilité d'implémenter des routines spécifiques à un terminal donné. L'architecture J2ME se découpe donc en plusieurs couches :

1. **Les profils** : Ils permettent à certaines catégories de terminaux d'utiliser des caractéristiques communes telles que la gestion de l'affichage, des événements d'entrées/sorties (pointage, clavier,...) ou des mécanismes de persistance (Base de données légère intégrée). Ces profils sont soumis à des spécifications suivant le principe

du JCP (Java Community Process).

2. **Les configurations** : Elles définissent une plate-forme minimale en terme de services concernant un ou plusieurs profils donnés.
3. **Les machines virtuelles** : En fonction de la cible, la machine virtuelle pourra être allégée afin de consommer plus ou moins de ressources (KVM,CVM,...).
4. **Le système d'exploitation** : L'environnement doit s'adapter au système d'exploitation existant (Windows CE, Symbian, Palm OS,SavaJe,...).

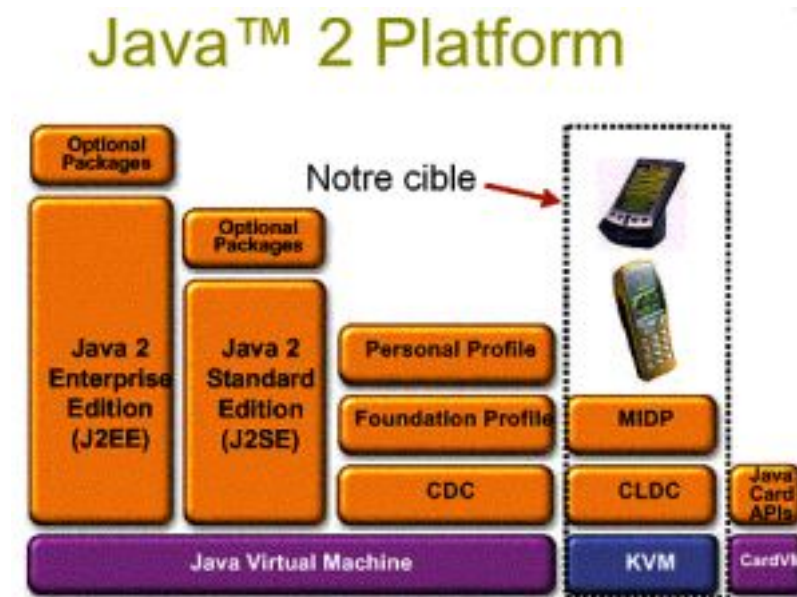


FIG. 7.1 – Java 2 plate-forme

Cette architecture en couche a pour but de factoriser pour des familles de produits données un ensemble d'API¹ permettant à une application de s'exécuter sur plusieurs terminaux sans modification de code. Dans cette optique, la plate-forme propose deux configurations :

- **CDC (Connected Device Configuration)** : Spécifie un environnement pour des terminaux connectés de forte capacité tels que les "set top boxes ", les téléphones à écran, la télévision numérique,...
- **CLDC (Connected Limited Device Configuration)** : Cible les périphériques à ressources limitées ou faibles tels que les téléphones mobiles, les assistants personnels, ou les périphériques légers sans fil (wireless)

¹Interface de programmation d'applications, contenant un ensemble de fonctions courantes de bas niveau, bien documentées, permettant de programmer des applications de "Haut Niveau".

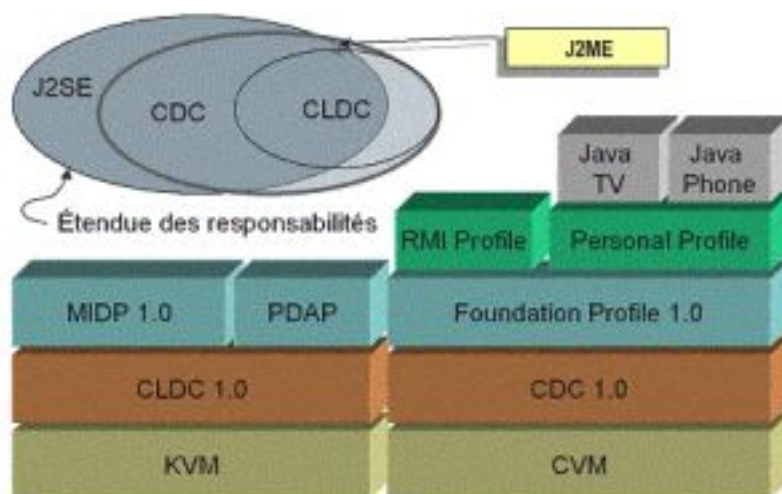


FIG. 7.2 – CDC et CLDC

CLDC (Connected Limited Device Configuration)

Voyons de plus près les contraintes en terme de ressources imposés par la configuration CLDC (celle qui nous intéresse). Cette configuration s'adresse aux terminaux légers tels que les téléphones mobiles ou les assistants personnels. Ces périphériques étant limités en terme de ressources, l'environnement classique ne permet pas de respecter les contraintes d'occupation mémoire liées à ces appareils. J2ME définit donc un ensemble d'API spécifiques à CLDC et destiné à utiliser les particularités de chaque terminal d'une telle famille (profil). La liste suivante résume l'ensemble de ces caractéristiques :

- Minimum de 160 à 512 Ko de RAM, processeur de 16 ou 32 bits, vitesse 16 Mhz ou plus.
- Alimentation limitée, prise en charge d'une batterie.
- Connexion réseau non permanente, sans fil.
- Interface graphique limitée ou inexistante définie par un sous-ensemble de classes Java s'exécutant dans le KVM (KiloByte Virtual Machine).

Cette configuration s'inscrit donc dans une logique d'économie de ressources avec une KVM de 40 à 80 Ko s'exécutant 30 à 80% moins vite qu'une JVM normale. Il n'y a aucun compilateur Just-in-Time ni même de prise en charge des nombres flottants. Quant au multi-threading et au Ramasse Miette, ils demeurent supportés. Toutefois CLDC n'intègre pas la gestion des interfaces graphiques, la persistance ou les particularités de chaque terminal. Ces aspects ne sont pas de sa responsabilité. Le tableau suivant reprend les packages et classes présentes dans cette couche :

Java.io	Fournit la gestion des flux d'entrées/sorties
Java.lang	Classes de base du langage (types,...)
Java.util	Contient les collections et classes utilitaires
Javax.microedition.io	Classes permettant de se connecter via TCP/IP

TAB. 7.1 – Liste des packages de CLDC

MIDP (Mobile Information Device Profile)

MIDP est à la base de l'implémentation des classes liées à un profil donné. On y trouve les méthodes permettant de gérer l'affichage, la saisie utilisateur et la gestion de la persistance (base de données). Il existe aujourd'hui deux implémentations majeures des profils MIDP. L'une plus spécifique destinée aux assistants de type Palm Pilot (Palm OS) et l'autre totalement générique proposée par Sun comme implémentation de référence (RI). Ces profils sont en libre téléchargement sur le site de Sun et intègrent plusieurs émulateurs permettant de tester les applications de manière logicielle. Les API liées à MIDP font aussi partie de ce package :

Javax.microedition.lcdui	Fournit la gestion de l'interface utilisateur (contrôles,...)
Java.microedition.midlet	Socle technique destiné à gérer le cycle de vie des midlets
Java.microedition.rms	Base de données persistante légère

TAB. 7.2 – Liste des packages de MIDP

L'API *lcdui* est chargée de gérer l'ensemble des contrôles graphiques proposés par ce profil. Quant à la gestion des événements, elle suit le modèle des **listeners** du J2SE avec un **CommandListener** appelé en cas d'activation d'un contrôle. Pour finir, *io* et *rms*, eux, fournissent les routines nécessaires aux entrées/sorties réseau et à la prise en charge d'une zone physique de stockage.

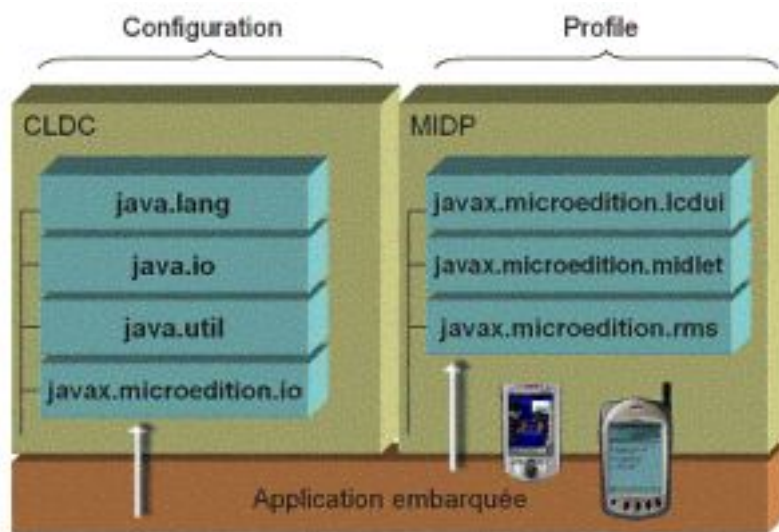


FIG. 7.3 – Les packages CLDC et MIDP

Les Midlets

Les Midlets sont l'élément principal d'une application Java embarquée. Pour bien saisir leur mode de fonctionnement, il suffit de prendre comme analogie les Applets ou les Servlets. Le cycle de vie d'un Applet est géré par un conteneur, en l'occurrence le navigateur Web, dont le rôle est d'interagir avec celle-ci sous la forme de méthodes de notifications prédéfinies (*init()*, *paint()*, *destroyed()*,...). Un Servlet possède les mêmes caractéristiques qu'un Applet excepté le fait que le conteneur est un moteur de servlet (Tomcat, WebSphere, WebLogic,...). Quant aux Midlets, ils représentent le pendant des Applets et des Servlets pour J2ME avec comme conteneur le téléphone mobile ou l'assistant personnel. Ainsi, en cas de mise à jour d'une application embarquée, un simple téléchargement de code Midlet est nécessaire à partir d'un quelconque serveur. De cette manière, un programme développé pour un profil donné est en mesure de s'exécuter sur tous les périphériques correspondant à cette famille. C'est aussi une manière de découpler le socle technique des applicatifs puisque le seul lien existant entre les logiciels embarqués et le terminal est l'API J2ME.

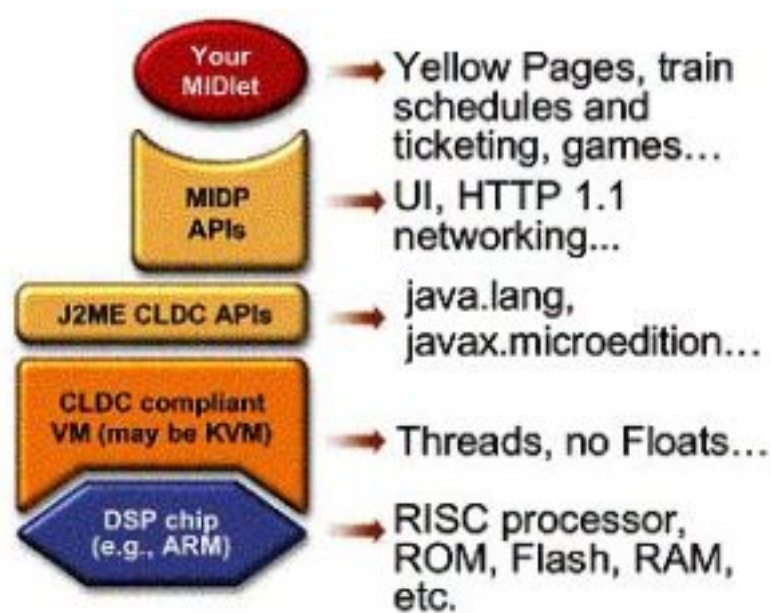


FIG. 7.4 – Les Midlets

7.2 Démarche

7.2.1 Introduction

Dans cette partie, nous allons voir comment l'application a été réalisée. Comment faut-il faire pour récupérer les informations boursières nécessaires pour l'application ? Où les calculs des centres mobiles doivent-ils être faits ? Le GSM doit-il faire des calculs ou bien n'affiche-t-il que les résultats ? Ces trois questions ont été au centre de la plupart des mes réflexions lors de l'implémentation surtout à cause de deux principales limites des téléphones portables à savoir la mémoire et la vitesse du processeur. La taille de l'écran est plus un problème lors de la création des différentes visualisations.

7.2.2 Architecture de l'application

Le calcul des K-Means Le calcul des K-Means étant un calcul demandant un processeur rapide, le mettre sur le téléphone mobile ne s'avère pas une très bonne idée. Par contre, construire un serveur Web qui téléchargerait les données voulues pour ensuite calculer sur ces données chaque cluster et leur valeur et finalement renvoyer les résultats semble une bien meilleure idée dans le sens où le calcul des K-Means tournerait sur la machine hébergeant le serveur, le processeur de cette machine est très certainement supérieure à la plupart des processeurs de GSM.

Les visualisations Par contre, les visualisations de l'application n'ont besoin que des résultats du calcul des K-Means plus quelques données pour l'évolution des clusters dans le cas du Pixel Bar Chart et de la Spirale. La construction des visualisations va donc se faire sur le GSM après récupération des résultats.

Connexion GSM-Serveur Finalement, l'application a été développée avec Sun One Studio 4 qui offre des facilités pour connecter les mobiles avec des serveurs en construisant automatiquement deux classes, une sur le GSM et une sur le serveur qui permettent à ceux-ci de communiquer directement. Cependant, l'application tourne bien avec l'émulateur mais, par manque de temps, les tests ont été réalisés en **local** avec des données pré-enregistrées sur le GSM. Il faut rappeler que le but principal des tests était avant tout de tester les différentes visualisations.

7.2.3 Récupération des données boursières

La technique utilisée pour récupérer les données est le téléchargement sur Internet. Plusieurs solutions sont possibles pour télécharger les données comme les Web Services. Mais ceux-ci par exemple sont payants et comme une autre solution gratuite existe, notre choix s'est basé sur celle-ci. Il existe quand même des Web Services gratuits mais tous ceux rencontrés n'offrent que très peu de services (en générale les valeurs boursières du jour). Ainsi, la solution adoptée est le téléchargement sur les sites boursiers car ceux-ci sont gratuits, sont permanents et sont sans cesse mis à jour. Pour cela, un parseur existe sur notre serveur qui va parser les pages Web du site boursier (yahoo Finance est le site utilisé par l'application) contenant les données voulues.

L'adresse html étant connue d'avance, il suffit de modifier quelques variables dans cette adresse pour obtenir l'historique de n'importe quelle action sur n'importe quel intervalle. Dans Java, l'adresse URL étant créée par le constructeur "new URL(String URL)" les variables de l'URL du site yahoo sont donc des "String" qui sont initialisés suivant l'intervalle de temps que veut l'utilisateur, ensuite ils sont concaténés à l'URL au bon endroit. Les actions de l'utilisateur (de son portefeuille) sont quant à elles contenues dans une liste de "String", chacune sous la forme de leur symbole boursier car seul le symbole est requis dans l'URL pour avoir les informations sur l'action.

7.2.4 Le calcul des K-Means

Le calcul des K-Means est très important car il est à la base de l'application. Pour réaliser ce calcul, l'utilisation de Thread² a été requise.

²Processus léger, correspondant à l'exécution d'un petit programme, ou d'une routine d'un programme plus gros, indépendamment de celui-ci.

L'ensemble des données va d'abord être découpé entre les différents Threads. Le nombre de Thread pour le calcul est facultatif et laissé au soin du programmeur qui peut choisir si 2 Threads travailleront ou bien 5. Les données vont être regroupées dans un tableau (avec 2 lignes, le symbole de l'action sur la première et sa valeur sur la deuxième) et les différents Threads s'occuperont chacun au début d'une partie du tableau. Le premier Thread traitera par exemple les données de la colonne 1 à 30 du tableau et ainsi de suite pour les autres. Le nombre de données qu'un Thread doit traiter est fonction du nombre de colonnes du tableau des données et du nombre de Threads. Les centres vont quant à eux être initialisés aux K premières données (K étant donné par l'utilisateur). Ils vont être, comme les données, mis dans un tableau et être assignés à un Thread. Dans le cas où il y a plus de Thread que de centres (K est plus petit que le nombre de Thread), les K premiers Threads sont responsables des centres. Chaque Thread va donc être responsable d'un segment de tableau dans le tableau des données et d'un segment dans le tableau des centres.

Chaque Thread va ensuite posséder des statistiques personnelles c'est à dire un tableau contenant la somme des valeurs des points appartenant à un cluster et ce, pour tous les clusters et un tableau contenant le nombre de points qui se trouvent dans chaque cluster.

Tous les Threads vont ensuite être démarrés. Ils vont rentrer dans une boucle et n'en sortiront qu'à la condition d'arrêt. Cette condition est qu'aucun centre n'a changé entre deux itérations de cette boucle.

Dans la boucle, chaque Thread va calculer les statistiques des centres. Il va pour cela calculer pour chacun des points contenus dans son segment de données, le centre le plus proche et changer à chaque fois ses statistiques personnelles. Quand chaque point de son segment aura été traité, il va mettre à jour les statistiques générales appartenant à tous les Threads (un tableau accumulateur général comme celui pour chaque Thread et un tableau général du nombre de points dans chaque cluster) pour nettoyer ensuite ses propres statistiques et les remettre à zéro.

Pour calculer le centre le plus près d'un point, la méthode est très simple. Le Thread calcule dans le tableau des centres celui qui est le plus près en calculant la distance euclidienne. Il va parcourir le tableau en calculant pour tous les centres, leur distance par rapport au point pour ne retenir que le centre minimum (grâce à son index dans le tableau).

Quand le calcul des statistiques des centres est fini, le Thread va mettre à jour ses centres c'est à dire les centres pour lesquels il est responsable. Cependant comment faire

pour synchroniser tous les Threads ? Car si un thread met à jour ses centres et qu'un autre n'a pas fini son calcul et modifie toujours des centres, il pourrait y avoir des erreurs. La solution est d'utiliser une barrière (définie par une classe Java) pour les synchroniser tous. La barrière sera de la taille du nombre de Threads participant au calcul. A chaque fois qu'un Thread finit son calcul de statistique, il rentre dans la barrière et attend que celle-ci soit occupée par tous les Threads. Ensuite tous les Threads sont réveillés. La mise à jour des centres peut alors s'effectuer en toute sécurité. Pour mettre à jour ses centres, un Thread va utiliser l'accumulateur général (qui est un tableau) et divisera la valeur accumulée d'un centre (qui se trouve dans ce tableau à l'index du centre) pour ensuite le diviser par le nombre de points correspondant à ce centre qu'il peut trouver dans le tableau général contenant le nombre de points associés à chaque cluster.

Quand la mise à jour est effectuée, si aucun centre n'a été modifié, la condition d'arrêt est présente et le calcul s'achève. Les centres sont alors renvoyés comme résultats.

L'algorithme mémorise aussi quelle action est contenue dans quel cluster. Une variable permet d'enregistrer le contenu de chaque cluster au fur et à mesure que le calcul des K-Means se fait.

7.2.5 Implémentation des différentes visualisations

Nous ne rentrerons pas en détail dans l'implémentation des différentes visualisations car celle-ci n'est pas très compliquée. Malgré tout, pour une meilleure compréhension du code des différentes visualisations présent en Annexe, une présentation de l'interface utilisateur de bas niveau de J2ME est nécessaire.

L'interface utilisateur de bas niveau de J2ME comprend les classes **Canvas**, **Graphics** et **Font**. La classe **Canvas** permet d'écrire des applications pouvant accéder aux événements de bas niveau, offrant ainsi un grand contrôle sur l'affichage. Les jeux sont la meilleure illustration du type d'application qui utilise ce mécanisme. Elle comprend également la classe **Graphics**. Celle-ci permet de produire des graphiques en 2D. Elle est similaire à la classe **java.awt.Graphics** de J2SE. C'est cette classe qui nous permet de créer les différentes visualisations grâce à la méthode *protected void paint(Graphics g)* qu'elle contient. La classe **Font** représente les polices de caractères ainsi que les métriques associées.

Rappelons que les réels ne sont pas supportés par J2ME et donc il faut recourir à une transformation pour pouvoir les utiliser. Lorsque le GSM reçoit les résultats, il les reçoit sous forme de "String". Pour supprimer la virgule d'une valeur, les chiffres de cette

valeur sont parsés à partir de la gauche jusqu'à la virgule. Les chiffres de gauche sont ensuite concaténés à ceux de droite (qui sont les chiffres décimaux) sans la virgule. De cette façon, les valeurs sont ainsi converties en entiers et elles peuvent être manipulées à présent. Pour revenir à des chiffres avec virgules pour les afficher à l'écran par exemple, il suffit de réinsérer la virgule au bon endroit mais pour cela, nous posons la condition que les valeurs utilisées dans cette application depuis le site boursier jusqu'au GSM n'ont que 2 décimales. Dès lors, nous pouvons sans craindre remettre la virgule à la bonne place.

7.3 Interface

L'application commence tout d'abord par un menu principal dans lequel l'utilisateur peut choisir entre 3 choix. Le premier est l'application en elle-même (choix=Appllication), le second est un texte expliquant la technique du clustering (choix=Clustering) et le troisième est un texte sur la réalisation de l'application (choix=About) c'est à dire que celle-ci a été réalisé dans but d'un mémoire,... La figure 7.5 représente ce menu.



FIG. 7.5 – Menu principal de l'application

Lorsque l'utilisateur a choisi l'application, il doit faire un choix de période c'est à dire qu'il va choisir entre l'intervalle de 1 mois ou celui de 3 mois (figure 7.6). Ensuite, le choix du domaine des valeurs s'impose à lui. Il doit choisir entre les cours des actions ou bien les variations des cours des actions (figure 7.7).



FIG. 7.6 – Choix de la période (choix exclusif).



FIG. 7.7 – Choix du domaine des valeurs (choix exclusif).

Après avoir choisi la période et le domaine des valeurs, l'utilisateur doit donner le nombre de clusters (classes) qu'il souhaite voir dans le résultat (figure 7.8).



FIG. 7.8 – Choix du nombre de clusters.

A présent que toutes les variables ont été choisies par l'utilisateur, le menu des visualisations apparaît (figure 7.9). L'utilisateur peut aller au gré voir chaque visualisation. Celle-ci ne seront pas présentées ici car elle l'ont été dans le chapitre 6. Cependant pour la Spirale, l'utilisateur doit préciser quels clusters il veut comparer (figure 7.10). Si l'utilisateur donne un cluster de numéro supérieur au nombre de clusters qu'il avait choisi au préalable, une erreur apparaît à l'écran pendant quelques secondes et l'utilisateur doit ensuite changer de numéro (figure 7.11).



FIG. 7.9 – Menu des visualisations.



FIG. 7.10 – Choix du premier cluster représentant la spirale extérieure. Le choix du deuxième suit celui-ci et est le même.



FIG. 7.11 – Texte apparaissant à l'écran pendant quelques secondes si le numéro du cluster choisi est supérieur au nombre de clusters choisi par l'utilisateur au préalable.

7.4 Problèmes et Solutions

7.4.1 Lenteur lors de la récupération des données

La récupération des données est très lente car le serveur doit parser des pages Web, ce qui prend beaucoup de temps. C'est pourquoi l'application ne présente que des intervalles courts de 1 mois ou 3 mois. Les intervalles de 6 mois ou un an ne sont pas envisageables car l'application prendrait trop de temps pour télécharger les données. L'attente de la récupération des données est de l'ordre d'une trentaine de secondes pour l'intervalle de 3 mois car les pages Web sont longues et donc le parseur prend plus de temps à les parcourir.

Il existe plusieurs solutions. D'abord, l'intervalle de temps pourrait encore être réduit et pourrait être de l'ordre d'une semaine ou de quelques jours. En effet, les cours des actions qui intéressent le plus en général sont ceux du jour même ou bien d'une semaine auparavant au maximum. Il n'est donc pas nécessaire d'aller si loin dans le temps pour donner des informations. Ensuite, il existe certainement d'autres techniques (les Web

services par exemple) qui permettent de récupérer des informations boursières même si ceux-ci sont payants.

7.4.2 Lenteur lors de l’affichage de certaines visualisations

Le Pixel Bar Chart présente des faiblesses d’affichage. La lenteur de son affichage (3 à 5 secondes) est due au fait que c’est le GSM qui construit cette visualisation. La lenteur du processeur est donc en cause. La spirale prend aussi du temps à s’afficher mais moins que le Pixel Bar Chart car elle possède moins de couleurs à afficher que celui-ci. Le Bar Chart et le tableau, quant à eux, n’ont pas de problèmes d’affichage.

Une solution à ce problème serait de laisser le soin au serveur de construire les différentes visualisations après avoir effectué le calcul des K-Means pour ensuite renvoyer ces visualisations sous forme d’images au GSM. Le GSM peut ensuite afficher ces images normalement comme toute autre image ce qui ne prend guère de temps.

Chapitre 8

Expérience

8.1 Introduction

Nous allons dans cette section parcourir le questionnaire posé aux utilisateur pour ensuite évaluer l'application grâce à des tests utilisateurs qui ont été fait sur l'application avec une vingtaine d'utilisateurs.

8.2 Evaluation

8.2.1 L'échantillon

Au total, l'échantillon contient 20 personnes. Voyons de plus près sa composition suivant le sexe, l'âge, la profession, l'expérience en économie, l'expérience en informatique et l'expérience pour les téléphones mobiles :

- Sexe : l'échantillon comprend 4 femmes et 16 hommes.
- Age :
 - Entre 20 et 24 ans : 15 personnes.
 - Entre 25 et 39 ans : 3 personnes.
 - 40 ans et + : 2 personnes.
- Profession :
 - 10 étudiants en informatique.
 - 1 professeur d'économie.
 - 1 professeur d'informatique de gestion.
 - 1 gérant d'une station essence.
 - 1 retraité.
 - 2 étudiants en ingénieur civil.
 - 2 étudiants en comptabilité.
 - 1 employé au théâtre de Namur.

- 1 employé en graphisme (publicité).
- Expérience en économie :
 - Moins de 1 ans : 6 personnes.
 - Entre 1 et 3 ans : 9 personnes.
 - 4 ans et + : 5 personnes.
- Expérience en informatique :
 - Moins de 5 ans : 5 personnes.
 - Entre 5 et 9 ans : 7 personnes.
 - 8 ans et + : 8 personnes.
- Expérience pour les téléphones mobiles :
 - Moins de 3 ans : 3 personnes.
 - Entre 3 et 5 ans : 6 personnes.
 - 6 ans et + : 11 personnes.

8.2.2 Présentation du questionnaire

Informations personnelles

Tout d'abord, les volontaires devaient, avant de commencer les tests proprement dit, donner quelques informations personnelles.

Nom	
Prénom	
Profession	
Domaine de la profession	
Age	
Sexe	

TAB. 8.1 – Informations personnelles.

Expérience

Ensuite, il leur est demandé de noter leurs expériences dans les domaines tels que l'informatique, la bourse, les GSM,...

Expérience dans le domaine de l'informatique	
Expérience dans les domaines liés à la bourse	
Expérience dans les "Mobiles Devices"(PDA, GSM,...)	
Expérience dans les GSM	

TAB. 8.2 – Expérience.

Ensuite, une explication est donnée aux utilisateurs pour leurs permettre de naviguer dans l'application et de comprendre chacune des différentes visualisations. Les utilisateurs peuvent, après cela, commencer les différents scénarios.

Les Scénarios

Les utilisateurs vont enfin pouvoir répondre à une série de questions divisées en scénarios suivant chaque visualisation. Les questions posées sont en fonction des scénarios et en général aucune erreur n'a été détectée dans les réponses des utilisateurs excepté quelques unes qui sont plutôt du à des erreurs de distraction. Ceci s'explique par le fait que les questions posées étaient très simples (genre trouver le minimum d'un Bar Chart). Dès lors aucune mesure de temps de réponse des utilisateurs n'a été prise car la plupart des utilisateurs prenaient le même temps et en plus répondaient très vite aux questions posées. Passons à présent à la présentation des différents scénarios.

Scénario 1 Le premier scénario consiste à utiliser le cours des actions, une période de trois mois, sept clusters et comme visualisation le Bar Chart. Il doit ensuite trouver le maximum (et sa valeur) du Bar Chart, le minimum (et sa valeur) du Bar Chart et le contenu de certains clusters donnés.

Scénario 2 Ce scénario est identique au précédent sauf que l'utilisateur va travailler sur les variations journalières des actions. Des nombres négatifs apparaîtront donc dans les résultats. L'intervalle de temps est changé aussi et devient un mois. Les questions posées à l'utilisateur sont du même ordre que dans le premier scénario mais en plus il doit dire combien il y a de positifs et négatifs dans le Bar Chart.

Scénario 3 Le scénario est identique au scénario 2 excepté le fait que la visualisation sera le Bar Chart de type 2 que l'utilisateur pourra choisir en changeant les options. L'intervalle est par contre ici de 3 mois.

Scénario 4 Le troisième scénario concerne le tableau et pose les mêmes questions que les trois autres scénarios à part que l'on ne demande pas à l'utilisateur le contenu d'un cluster car cette visualisation ne le donne pas.

Scénario 5 Le Pixel Bar Chart est testé par ce scénario. L'intervalle est de un mois et le nombre de clusters est de 7. Les questions sont du même type que les autres scénarios mais l'utilisateur doit en plus répondre à des questions sur l'évolution de certains clusters.

Scénario 6 Ce scénario est différent des autres car il utilise la Spirale qui compare l'évolution de deux clusters. Les questions sont assez simples et demandent à l'utilisateur de comparer approximativement l'évolution de deux clusters dans le temps.

Scénario 7 Finalement, des questions sur l'utilisation du clustering pour le GSM sont posées ainsi que toute remarque sur les différentes visualisations.

8.2.3 Utilisabilité

L'utilisabilité est un concept très important pour les téléphones mobiles. C'est pourquoi après avoir fait les tests sur les différentes visualisations de l'application, les utilisateurs ont dû répondre à une série de questions sur l'utilisabilité de l'application. Celle-ci comporte-t-elle des bugs ? La navigation est-elle intuitive ? L'utilisateur pouvait aussi ajouter ses propres remarques sur l'application.

8.2.4 Résultats

Introduction

Les tests ont été passés sur un GSM Nokia 6100 qui est un GSM basé sur la technologie Java. Les préférences des utilisateurs ont été mis dans un tableau (tableau 8.3) avec quatre valeurs qualitatives : Mauvais, Moyen, Bon, Excellent.

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

TAB. 8.3 – Tableau des différentes valeurs qualitatives.

Test du Chi-carré

Le but du test est de voir si les préférences varient d'une représentation à l'autre. Nous avons pour cela choisi d'utiliser la méthode du Chi-Carré. Pour faire ce test, nous allons regrouper les préférences en deux classes pour ne pas avoir de nombres nuls dans le tableau des fréquences (le tableau 8.4 représente les fréquences initiales observées). La classe 1 va regrouper les préférences Mauvais et Moyen, et la classe 2 les préférences Bon et Excellent. Nous allons à présent tester s'il y a des préférences des utilisateurs pour certaines visualisations ou pas. Donc notre hypothèse H_0 sera que les réponses des

utilisateurs sont semblables quelle que soit la représentation. Le tableau 8.5 reprend les différentes fréquences (f_{ij}) des deux classes.

f_{ij}	Mauvais	Moyen	Bon	Excellent	f_i
Bar Chart 1	0	10	10	0	20
Bar Chart 2	4	6	8	2	20
Tableau	0	14	6	0	20
Pixel Bar Chart	7	3	7	3	20
Spirale	12	7	1	0	20
f_j	63	37	100 = N		

TAB. 8.4 – Tableau des fréquences initiales observées.

f_{ij}	Mauvais/Moyen	Bon/Excellent	f_i
Bar Chart 1	10	10	20
Bar Chart 2	10	10	20
Tableau	14	6	20
Pixel Bar Chart	10	10	20
Spirale	19	1	20
f_j	63	37	100 = N

TAB. 8.5 – Tableau des fréquences observées.

Les tableaux 8.4 et 8.5 montrent que l'avis est plus négatif pour le Tableau et la Spirale que pour les autres. Voyons si ce résultat est significatif. Il faut pour cela calculer les fréquences théoriques \tilde{f}_{ij} tel que $\tilde{f}_{ij} = \frac{f_j}{N} \times f_i$. Le tableau 8.6 représente le tableau des fréquences théoriques.

\tilde{f}_{ij}	Classe 1	Classe 2	\tilde{f}_i
Bar Chart 1	$\frac{63}{100} \times 20 = 12,6$	$\frac{37}{100} \times 20 = 7,4$	20
Bar Chart 2	$\frac{63}{100} \times 20 = 12,6$	$\frac{37}{100} \times 20 = 7,4$	20
Tableau	$\frac{63}{100} \times 20 = 12,6$	$\frac{37}{100} \times 20 = 7,4$	20
Pixel Bar Chart	$\frac{63}{100} \times 20 = 12,6$	$\frac{37}{100} \times 20 = 7,4$	20
Spirale	$\frac{63}{100} \times 20 = 12,6$	$\frac{37}{100} \times 20 = 7,4$	20
\tilde{f}_j	63	37	100

TAB. 8.6 – Tableau des fréquences théoriques.

Pour calculer χ^2 observé, il suffit d'utiliser la formule :

$$\chi_{obs}^2 = \sum_i \sum_j \frac{(f_{ij} - \tilde{f}_{ij})^2}{\tilde{f}_{ij}}$$

Dès lors on obtient le tableau 8.7.

	Classe 1	Classe 2	Total
Bar Chart 1	$\frac{(10-12,6)^2}{12,6} = 0,536$	$\frac{(10-7,4)^2}{7,4} = 0,913$	1,449
Bar Chart 2	$\frac{(10-12,6)^2}{12,6} = 0,536$	$\frac{(10-7,4)^2}{7,4} = 0,913$	1,449
Tableau	$\frac{(14-12,6)^2}{12,6} = 0,155$	$\frac{(6-7,4)^2}{7,4} = 0,264$	0,419
Pixel Bar Chart	$\frac{(10-12,6)^2}{12,6} = 0,536$	$\frac{(10-7,4)^2}{7,4} = 0,913$	1,449
Spirale	$\frac{(19-12,6)^2}{12,6} = 3,250$	$\frac{(1-7,4)^2}{7,4} = 5,535$	8,785
Total	5,013	8,538	13,551 = χ_{obs}^2

TAB. 8.7 – Calcul du χ_{obs}^2 .

Le nombre de degré de liberté associé à la statistique χ^2 est :

$$ddl = (l - 1) \times (c - 1) = 4 \times 1 = 4$$

où l est le nombre de lignes du tableau et c est le nombre de colonnes.

Grâce au degré de liberté, nous pouvons aller voir dans la Table de la loi du χ^2 pour trouver la valeur du χ^2 théorique en prenant $\alpha = 0,05$. On trouve donc la valeur du χ^2 théorique : $\chi_{0,05}^2 = 9,488$

L'hypothèse H_0 est donc fausse car $\chi_{theo}^2 < \chi_{obs}^2$. Les utilisateurs ont donc un avis différent pour certaines visualisations. Calculons à présent le coefficient de Cramer qui est une forme centrée réduite du Khi-carré :

$$K = \sqrt{\frac{\chi^2}{N - \min(l - 1, c - 1)}} = \sqrt{\frac{13,551}{20 \times 1}} = 0,823$$

où l est le nombre de lignes du tableau et c est le nombre de colonnes.

Le coefficient de Cramer varie entre 0 et 1 et plus la valeur de K est proche de 1, plus il y a corrélation. Ici, le coefficient de Cramer est plus près de 1 que de 0, donc on peut en conclure que les préférences des utilisateurs dépendent bien des différentes visualisations et changent suivant celles-ci.

Calcul des Médianes des différentes visualisations

Nous allons à présent calculer la médiane pour chaque visualisation. Elle permet de définir la tendance centrale d'une série en déterminant la valeur de l'individu qui se situe exactement à mi-classement[Mouchart]. La médiane est donc la valeur qui sépare la population en deux classes d'effectifs égaux. Ce paramètre est utile pour donner la

répartition du caractère étudié, car 50% environ de la population étudiée a une modalité inférieure à la médiane et 50% une modalité supérieure à la médiane.

Commençons d'abord par le Bar Chart de type 1. Le tableau 8.8 reprend les différents effectifs du Bar Chart 1 de type ainsi que les effectifs cumulés. On peut voir dans ce tableau que la médiane se trouve entre la préférence Moyen et la préférence Bonne. Si nous remplaçons les préférences par des valeurs ordinales (voir tableau 8.13), la médiane du Bar Chart de type 1 est de 2.5.

Préférence	Effectif	Effectif cumulé
Mauvais	0	0
Moyen	10	10
Bonne	10	20
Excellent	0	0

TAB. 8.8 – Tableau des effectifs et des effectifs cumulés du Bar Chart de type 1.

Passons ensuite au Bar Chart de type 2. Le tableau 8.9 reprend les différents effectifs du Bar Chart de type 2 ainsi que les effectifs cumulés. Nous pouvons voir dans ce tableau que la médiane se trouve entre la préférence Moyen et la préférence Bonne. Grâce au calcul de la médiane nous pouvons aussi remarquer que la médiane du Bar Chart de type 1 est la même que le Bar Chart de type 2. De même en remplaçant par les valeurs ordinales du tableau 8.13, la médiane du Bar Chart de type 2 est de 2.5.

Préférence	Effectif	Effectif cumulé
Mauvais	4	4
Moyen	6	10
Bonne	8	18
Excellent	2	20

TAB. 8.9 – Tableau des effectifs et des effectifs cumulés du Bar Chart de type 2.

Maintenant calculons la médiane du Tableau. Le tableau 8.10 reprend les différents effectifs du Tableau ainsi que les effectifs cumulés. La médiane ici se situe dans la préférence Moyen. En utilisant les valeurs ordinales du tableau 8.13, la médiane du Tableau est de 2.

Préférence	Effectif	Effectif cumulé
Mauvais	0	0
Moyen	14	14
Bonne	6	20
Excellent	0	0

TAB. 8.10 – Tableau des effectifs et des effectifs cumulés du Tableau.

Ensuite passons au Pixel Bar Chart. Le tableau 8.11 reprend les différents effectifs du Pixel Bar Chart ainsi que les effectifs cumulés. Nous pouvons voir dans ce tableau que la médiane se trouve encore entre la préférence Moyen et la préférence Bonne. De même en remplaçant par les valeurs ordinales du tableau 8.13, la médiane du Pixel Bar Chart est de 2.5.

Préférence	Effectif	Effectif cumulé
Mauvais	7	7
Moyen	3	10
Bonne	7	17
Excellent	3	20

TAB. 8.11 – Tableau des effectifs et des effectifs cumulés du Pixel Bar Chart.

Finalement calculons la médiane de la Spirale. Le tableau 8.12 reprend les différents effectifs de la Spirale ainsi que les effectifs cumulés. La médiane ici se situe dans la préférence Mauvais. En utilisant les valeurs ordinales du tableau 8.13, la médiane de la Spirale est de 1.

Préférence	Effectif	Effectif cumulé
Mauvais	12	12
Moyen	7	19
Bonne	1	20
Excellent	0	0

TAB. 8.12 – Tableau des effectifs et des effectifs cumulés de la Spirale.

Influence du Background

Pour analyser l'influence des différents facteurs du background des utilisateurs sur les préférences des différentes visualisations, nous avons utilisé l'environnement de data mining S-Plus [S-Plus]. Après avoir converti toutes nos données en un fichier excel, nous avons procédé à un **clustering** pour déterminer quels facteurs influencent le choix des visualisations malgré que le nombre d'observations est fort court pour cette technique. Les

facteurs sont au nombre de 4, l'expérience en informatique (en nombre d'années), l'expérience en économie (en nombre d'années), l'expérience pour les téléphones portables (en nombre d'années que l'utilisateur possède un GSM) et le sexe (Masculin ou féminin). Ensuite, les réponses qualitatives sur les différentes visualisations ont été transformées en valeurs ordinales comme montré dans le tableau 8.12.

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent
Valeur ordinale	1	2	3	4

TAB. 8.13 – Tableau des valeurs ordinales et des préférences correspondantes.

Pour rappel, le clustering est un procédé pour grouper des objets physiques ou abstrait en classes d'objets similaires. Son principe est de minimiser la distance intraclasse et de maximiser la distance interclasse.

Après avoir identifié les deux classes principales de chaque visualisation, nous allons tester la différence des moyennes entre les deux clusters les plus différents pour voir si le résultat est significatif (c'est à dire que nous testons si les différents groupes obtenus sont significativement différents du point de vue des avis des utilisateurs sur les visualisations). Pour cela, nous allons effectuer un certain nombre de calculs pour obtenir l'intervalle de confiance de la différence des moyennes [Wonnacott & Wonnacott91] :

- Calculer le nombre de degré de liberté ddl suivant :

$$ddl = (n_1 - 1) + (n_2 - 1)$$

où n_1 est le nombre d'éléments dans la première classe et n_2 le nombre d'éléments dans la deuxième.

- Calculer s_p ensuite qui est la variance commune aux deux classes suivant la formule :

$$s_p^2 = (\sum (X_1 - \bar{X}_1)^2 + \sum (X_2 - \bar{X}_2)^2) / ddl$$

où X_i est une valeur de la classe i et \bar{X}_i est la moyenne de la classe i .

- Calculer l'intervalle de confiance à 95% ($\alpha = 0,050$) utilisant la formule basée sur la différence des moyennes sous la t distribution de Student (à cause de la petite taille de l'échantillon) :

$$\mu_1 - \mu_2 = (\bar{X}_1 - \bar{X}_2) \pm t_{0,25 \cdot s_p} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Une fois ceci fait, nous pouvons calculer la probabilité critique pc par :

$$pc = \frac{(\bar{X}_1 - \bar{X}_2)}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Utilisant la ddl_{th} ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous pouvons déterminer la valeur t_p pour laquelle $t > t_p$ avec $0 < .p < 1$.

Nous pourrions alors conclure que $pc < .p$. Dans notre cas, nous avons décidé que notre résultat est significatif si $pc < 0,05$.

Nous allons à présent parcourir chaque visualisation et voir les préférences de chaque utilisateur pour celle-ci en fonction de certains facteurs.

Bar Chart de type 1 Pour expliquer les préférences des utilisateurs pour le Bar Chart de type 1, le premier facteur à prendre en compte est le nombre d'années d'expérience en économie. On peut voir que 6 personnes ayant une expérience en économie d'au maximum 0.75 ans trouvent le Bar Chart de type 1 Moyen.

Ensuite les personnes ayant une expérience en économie entre 0.75 ans et 7.5 ans jugent le Bar Chart de type 1 bon. Ce groupe est d'ailleurs majoritaire et contient 9 personnes. Ceux dont l'expérience en économie est au-dessus de 7.5 ans le qualifient de Moyen.

On peut donc en conclure que les utilisateurs ayant de bonnes connaissances en économie apprécient le Bar Chart de type 1 tandis que ceux qui ont peu de connaissances ainsi que ceux qui ont de très bonnes connaissances n'ont pas d'avis au sujet de cette visualisation et la jugent moyenne ou sans grand intérêt. Ainsi nous pouvons voir que les économistes de bas niveau et donc qui n'ont pas beaucoup de notions en statistique le considèrent moyen. Les économistes avec de très bonnes connaissances le connaissent parfaitement et donc le jugent eux aussi moyen car ils le trouvent peut-être trop banal. Notons aussi qu'en moyenne cette visualisation est jugée bonne.

Voici les résultats de S-Plus pour le Bar Chart de type 1 :

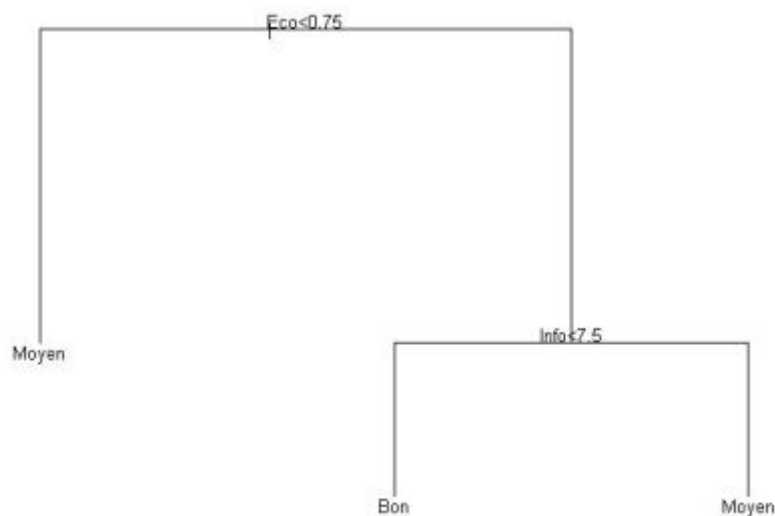
```

*** Tree Model ***

Classification tree:
tree(formula = Bar.Chart.1 ~ Info + Eco + GSM + Sexe, data = test, na.action
      = na.exclude, mincut = 5, minsize = 10, mindev = 0.01)
Variables actually used in tree construction:
[1] "Eco" "Info"
Number of terminal nodes: 3
Residual mean deviance: 1.275 = 21.67 / 17
Misclassification error rate: 0.25 = 5 / 20
node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 20 27.730 Bon ( 0.5000 0.5000 )
 2) Eco<0.75 6 5.407 Moyen ( 0.1667 0.8333 ) *
 3) Eco>0.75 14 18.250 Bon ( 0.6429 0.3571 )
   6) Info<7.5 9 9.535 Bon ( 0.7778 0.2222 ) *
   7) Info>7.5 5 6.730 Moyen ( 0.4000 0.6000 ) *
```

L'arbre de régression dérivée est :



Après avoir transformé les différentes préférences par des valeurs ordinales, calculons à présent la probabilité critique des deux principaux groupes à savoir ceux ayant une expérience en économie en-dessous de 0.75 ans et ceux au-dessus. Ici, $n_1 = 14$ et $n_2 = 6$. Dès lors :

- $ddl = (6 - 1) + (14 - 1) = 18$
- $s_p^2 = \frac{4,1812}{18} = 0,23$
- $(intervalle) = 0,48 \pm 2,101 \cdot \sqrt{0,23} \cdot \sqrt{\frac{1}{6} + \frac{1}{14}} = 0,48 \pm 0,239905762$
- $t = \frac{0,48}{\sqrt{0,23} \cdot \sqrt{\frac{1}{6} + \frac{1}{14}}} = 2.051171459$

Utilisant la 18th ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous obtenons $t > t_{.05}$ ce qui signifie que $p < .05$. Nos conclusions sont donc juste : l'expérience en économie a donc un impact sur la préférence pour le Bar Chart de type 1.

Bar Chart de type 2 Pour le Bar Chart de type 2, le premier facteur est l'expérience en informatique (en nombre d'années). Les utilisateurs dont l'expérience est inférieure à 4.5 ans en informatique (7 personnes) qualifient le Bar Chart de type 2 de bon.

Au-dessus de 4.5 ans en informatique, un deuxième facteur est à prendre en compte à savoir l'expérience en économie. Ceux qui ont une expérience au-dessus 2.5 ans le considère aussi Bon mais ceux en-dessous le trouve Mauvais cependant ceux-ci ne sont que 5.

En conclusion, nous pouvons dire que les utilisateurs ayant de bonnes connaissances en informatique et en économie préfèrent le Bar Chart de type 1 alors que les autres préfèrent le Bar Chart de type 2. Nous pouvons aussi dire que les personnes dont les notions en statistique sont bonnes (informaticiens et économistes accomplis) apprécient plus le Bar Chart de type 1 que le Bar Chart de type 2.

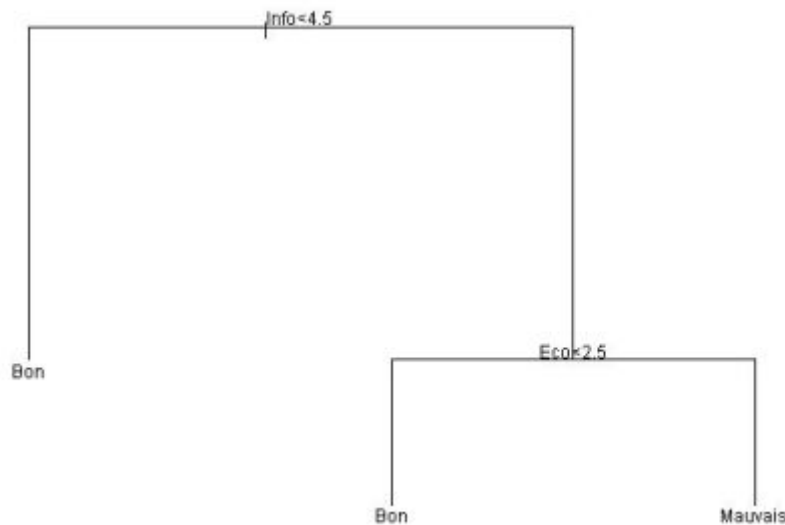
Voici les résultats de S-Plus pour le Bar Chart de type 2 :

```
*** Tree Model ***

Classification tree:
tree(formula = Bar.Chart.2 ~ Info + Eco + GSM + Sexe, data = test, na.action
      = na.exclude, mincut = 5, minsize = 10, mindev = 0.01)
Variables actually used in tree construction:
[1] "Info" "Eco"
Number of terminal nodes:  3
Residual mean deviance:  2.365 = 40.2 / 17
Misclassification error rate: 0.5 = 10 / 20
node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 20 51.190 Bon ( 0.4000 0.1000 0.2000 0.3000 )
 2) Info<4.5 7 15.110 Bon ( 0.4286 0.2857 0.0000 0.2857 ) *
 3) Info>4.5 13 28.410 Bon ( 0.3846 0.0000 0.3077 0.3077 )
    6) Eco<2.5 8 15.590 Bon ( 0.5000 0.0000 0.1250 0.3750 ) *
    7) Eco>2.5 5  9.503 Mauvais ( 0.2000 0.0000 0.6000 0.2000 ) *
```

L'arbre de régression dérivée :



Calculons à présent la probabilité critique des deux principaux groupes à savoir ceux ayant une expérience en informatique en-dessous de 4.5 ans et ceux au-dessus. Ici, $n_1 = 7$ et $n_2 = 13$. Dès lors :

- $ddl = (7 - 1) + (13 - 1) = 18$
- $s_p^2 = \frac{12,9237}{18} = 0,72$
- $(intervalle) = 0,93 \pm 2,101 \cdot \sqrt{0,72} \cdot \sqrt{\frac{1}{7} + \frac{1}{13}} = 0,93 \pm 0.39181486$
- $t = \frac{0,93}{\sqrt{0,72} \cdot \sqrt{\frac{1}{7} + \frac{1}{13}}} = 2.33788099$

Utilisant la 18th ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous obtenons $t > t_{.025}$ ce qui signifie que $p < .025$. Nos conclusions sont donc justes : l'expérience en informatique a donc un impact sur la préférence pour le Bar Chart de type 2.

Tableau Nous pouvons en général voir que le Tableau est une visualisation perçue moyennement par les utilisateurs. En effet, la plupart le considère comme Moyen. Cela est sûrement dû au fait que le Tableau est une visualisation courante et que les utilisateurs n'ont en général rien à dire à son sujet. Malgré tout, nous allons tenter de l'évaluer. On peut voir que le premier facteur influençant la préférence des utilisateurs est encore l'expérience en informatique. Comme déjà dit la plupart le juge Moyen mais certains utilisateurs ayant une expérience en informatique au-dessus de 4.5 ans (ou plutôt entre 4.5 et 6.5 ans) le qualifient de Bon.

Voici les résultats de S-Plus pour le Tableau :

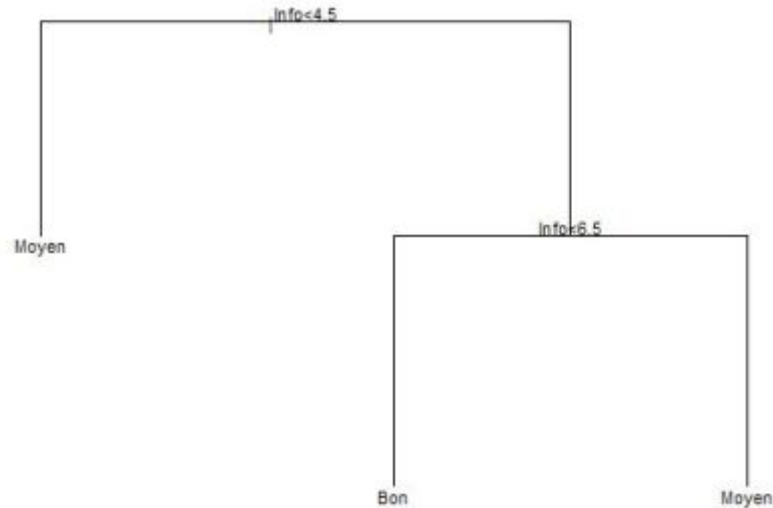
```

*** Tree Model ***

Classification tree:
tree(formula = Tableau ~ Info + Eco + GSM + Sexe, data = test, na.action =
      na.exclude, mincut = 5, minsize = 10, mindev = 0.01)
Variables actually used in tree construction:
[1] "Info"
Number of terminal nodes:  3
Residual mean deviance:  1.263 = 21.47 / 17
Misclassification error rate: 0.25 = 5 / 20
node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 20 24.430 Moyen ( 0.3000 0.7000 )
 2) Info<4.5 7  5.742 Moyen ( 0.1429 0.8571 ) *
 3) Info>4.5 13 17.320 Moyen ( 0.3846 0.6154 )
    6) Info<6.5 5  6.730 Bon ( 0.6000 0.4000 ) *
    7) Info>6.5 8  8.997 Moyen ( 0.2500 0.7500 ) *
```

L'arbre de régression dérivée :



Calculons la probabilité critique des deux principaux groupes à savoir ceux ayant une expérience en informatique en-dessous de 4.5 ans et ceux au-dessus. Ici, $n_1 = 7$ et $n_2 = 13$. Dès lors :

- $ddl = (7 - 1) + (13 - 1) = 18$
- $s_p^2 = \frac{3,9344}{18} = 0,22$
- $(intervalle) = 0,24 \pm 2,101 \cdot \sqrt{0,22} \cdot \sqrt{\frac{1}{7} + \frac{1}{13}} = 0,24 \pm 0.216583813$
- $t = \frac{0,24}{\sqrt{0,22} \cdot \sqrt{\frac{1}{7} + \frac{1}{13}}} = 1.091454411$

Utilisant la 18th ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous obtenons $t < t_{.05}$ ce qui signifie que $p > .05$. Notre résultat n'est donc pas juste : le Tableau n'est pas influencé par l'expérience en informatique. L'échantillon comportant beaucoup d'informaticiens, on peut en conclure que certains trouvent bon le Tableau d'autres le trouvent moyen mais cela n'a aucunement été influencé par leurs années d'expérience en informatique.

Pixel Bar Chart Nous évaluons à présent une des deux nouvelles visualisations. Nous pouvons commencer par voir que le nombre d'années d'expérience dans les téléphones mobiles est le premier facteur à considérer. Les utilisateurs avec 5.5 ans d'expérience jugent le Pixel Bar Chart de mauvais.

Ensuite, pour les utilisateurs dont l'expérience pour les GSM est au-dessus de 5.5 ans malgré que, en moyenne, leur préférence est bonne pour cette visualisation, il faut considérer un deuxième facteur à savoir le nombre d'années d'expérience en informatique. Lorsque celle-ci dépasse 6.5 ans, le Pixel Bar Chart est jugé de bon voir même Excellent pour certains utilisateurs (2 utilisateurs sur 5 le jugent d'excellent).

Nous pouvons dès lors voir que les utilisateurs qui ont au moins 6.5 ans d'expérience en informatique et donc ont eu une bonne formation en statistique ont un avis positif envers le Pixel Bar Chart. En conclusion, nous pouvons dire que plus les connaissances en statistique sont élevées, plus le Pixel Bar Chart est apprécié.

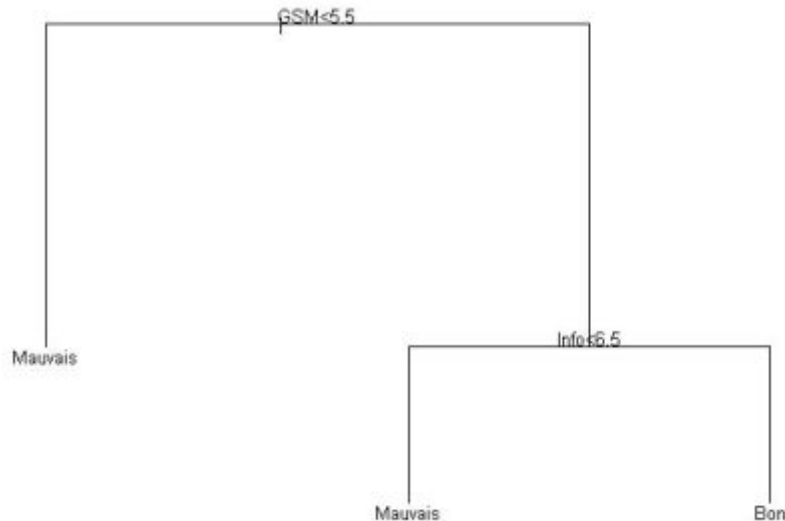
Voici les résultats de S-Plus pour le Pixel Bar Chart :

```
*** Tree Model ***

Classification tree:
tree(formula = Pixel ~ Info + Eco + GSM + Sexe, data = test, na.action =
      na.exclude, mincut = 5, minsize = 10, mindev = 0.01)
Variables actually used in tree construction:
[1] "GSM" "Info"
Number of terminal nodes:  3
Residual mean deviance:  2.233 = 37.96 / 17
Misclassification error rate: 0.5 = 10 / 20
node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 20 52.16 Bon ( 0.3500 0.1500 0.3500 0.1500 )
 2) GSM<5.5 9 19.10 Mauvais ( 0.2222 0.0000 0.4444 0.3333 ) *
 3) GSM>5.5 11 23.48 Bon ( 0.4545 0.2727 0.2727 0.0000 )
    6) Info<6.5 6 12.14 Mauvais ( 0.3333 0.1667 0.5000 0.0000 ) *
    7) Info>6.5 5  6.73 Bon ( 0.6000 0.4000 0.0000 0.0000 ) *
```

L'arbre de régression dérivée :



Calculons maintenant la probabilité critique des deux principaux groupes à savoir ceux ayant une expérience en informatique en-dessous de 4.5 ans et ceux au-dessus. Ici, $n_1 = 9$ et $n_2 = 11$. Dès lors :

- $ddl = (9 - 1) + (11 - 1) = 18$
- $s_p^2 = \frac{19,7334}{18} = 1,09$
- $(intervalle) = 0,95 \pm 2,101 \cdot \sqrt{1,09} \cdot \sqrt{\frac{1}{9} + \frac{1}{11}} = 0,95 \pm 0,443133009$
- $t = \frac{0,95}{\sqrt{1,09} \cdot \sqrt{\frac{1}{9} + \frac{1}{11}}} = 2,024477506$

Utilisant la 18th ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous obtenons $t > t_{.05}$ ce qui signifie que $p < .05$. Nos conclusions sont donc justes : le nombre d'années d'expérience pour les téléphones portables influence les préférences des utilisateurs pour le Pixel Bar Chart.

Spirale Lors des tests, cette visualisation a été la moins bien perçue par les utilisateurs et beaucoup d'explications à son sujet ont dû être fournies aux utilisateurs pour leur permettre de la comprendre. Par l'évaluation, on peut voir que les utilisateurs qualifient celle-ci de mauvaise même si certains la trouvent moyenne. On peut voir que le premier facteur influençant le choix des utilisateurs est leur expérience en économie.

Ensuite, dans ceux qui ont une expérience supérieure à 0.75 ans, le groupe est encore divisé suivant le facteur économique. Nous pouvons dès lors voir que uniquement les personnes dont l'expérience en économie est au-dessus de 2.5 ans et donc avec une formation plus importante en statistique jugent cette visualisation de moyenne.

En conclusion, nous pouvons dire que cette visualisation n'a pas bien été acceptée par les utilisateurs tant par sa compréhension qui n'est pas facile que par sa lisibilité. Certains utilisateurs trouvaient qu'il n'était pas facile de comparer les deux spirales et ont proposé de mettre deux lignes horizontales pour pouvoir mieux comparer. Ceci équivaut à déplier les deux spirales.

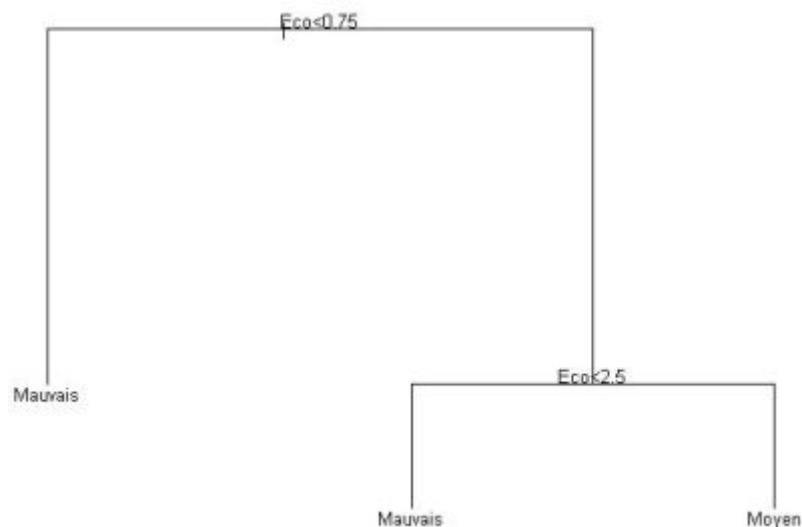
Voici les résultats de S-Plus pour la Spirale :

```
*** Tree Model ***

Classification tree:
tree(formula = Spirale ~ Info + Eco + GSM + Sexe, data = test, na.action =
      na.exclude, mincut = 5, minsize = 10, mindev = 0.01)
Variables actually used in tree construction:
[1] "Eco"
Number of terminal nodes:  3
Residual mean deviance:  1.667 = 28.35 / 17
Misclassification error rate: 0.35 = 7 / 20
node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 20 32.950 Mauvais ( 0.0500 0.6000 0.3500 )
 2) Eco<0.75 6 10.410 Mauvais ( 0.1667 0.6667 0.1667 ) *
 3) Eco>0.75 14 19.120 Mauvais ( 0.0000 0.5714 0.4286 )
    6) Eco<2.5 7  8.376 Mauvais ( 0.0000 0.7143 0.2857 ) *
    7) Eco>2.5 7  9.561 Moyen ( 0.0000 0.4286 0.5714 ) *
```

L'arbre de régression dérivée :



Calculons à présent la probabilité critique des deux principaux groupes à savoir ceux ayant une expérience en économie en-dessous de 0.75 ans et ceux au-dessus. Ici, $n_1 = 6$ et $n_2 = 14$. Dès lors :

- $ddl = (6 - 1) + (14 - 1) = 18$
- $s_p^2 = \frac{6,9296}{18} = 0,38$
- $(intervalle) = 0,08 \pm 2,101 \cdot \sqrt{0,38} \cdot \sqrt{\frac{1}{6} + \frac{1}{14}} = 0,08 \pm 0.308367471$
- $t = \frac{0,08}{\sqrt{0,38} \cdot \sqrt{\frac{1}{6} + \frac{1}{14}}} = 0,265963986$

Nos conclusions ne sont donc absolument pas justes : la préférence des utilisateurs pour la Spirale n'est pas influencée par l'expérience en économie car utilisant la 18th ligne de la table des valeurs critiques de la loi de Student disponible dans l'Appendice A, nous obtenons $t < t_{0.05}$ ce qui signifie que $p > .05$. De plus 0 fait partie de l'intervalle de confiance et donc aucune conclusion ne peut être faite.

La navigation La navigation n'a pas été évaluée comme les différentes visualisations car l'utilisateur devait critiquer celle-ci ou bien donner des conseils au lieu de la juger suivant une échelle de préférence. Evidemment, vu le peu de menus que contient l'application, il est normal que peu de critiques ont été fournies pas les différents utilisateurs. Néanmoins, nous pouvons nous apercevoir que deux principales critiques reviennent sans cesse :

- "Select-Ok" : Les tests ont, comme nous l'avons dit précédemment, été faits sur le GSM Nokia 6100. Cependant, l'application a été réalisée avec un émulateur et lorsque celle-ci a été mise sur le GSM, certains boutons ont changé de place (notamment le menu pour choisir l'intervalle de temps et le domaine des valeurs, où il n'y avait que deux boutons à savoir "Options" et "Back"). En appuyant sur "Options", un menu apparaissait laissant deux choix à l'utilisateur : "Select" et "Ok". L'utilisateur va utiliser "Select" pour choisir soit son intervalle de temps soit le domaine des valeurs qu'il veut suivant le menu dans lequel il est. Le "Ok" lui permettra de valider son choix et de passer au menu suivant. La plupart des utilisateurs ont trouvé ce principe très déroutant.
- Les "Back" : Quelques utilisateurs ont regretté qu'il n'y avait pas de boutons "Back" pour revenir au menu du début. Les boutons "Back" renvoient tous l'utilisateur au menu précédent mais aucun ne le renvoie au début de l'application.

8.3 Améliorations

8.3.1 La navigation

Vu la taille de l'application, peu d'améliorations peuvent lui être apportées. Malgré tout, le problème des boutons "Back" peut être résolu et offrir un bouton de retour au

menu principal. Le problème du "Select-Ok" est juste dû à l'utilisation spécifique du GSM Nokia 6100.

8.3.2 Les différentes visualisations

La première amélioration à apporter aux différentes visualisations est d'abord de réduire l'intervalle de temps sur lequel on prend les différentes valeurs. En général, les économistes sont plus intéressés par des historiques courts c'est à dire sur des intervalles d'au plus une semaine. Cette idée faciliterait l'implémentation des différentes visualisations car il y aurait moins de données à considérer. Le temps de calcul et de téléchargement des données en serait aussi réduit.

Une échelle de temps pour la Spirale et le Pixel Bar Chart pourrait aider l'utilisateur lorsqu'il veut voir l'évolution d'un cluster. De même, une échelle de couleur pourrait être aussi intéressante pour aider à voir l'évolution des clusters pour la Spirale et le Pixel Bar Chart.

La Spirale est la visualisation la moins appréciée par les utilisateurs. Une alternative à cette visualisation serait de déplier la Spirale et de mettre les évolutions sur des bandes horizontales deux par deux (car deux clusters à comparer) les unes en dessous des autres parce que la largeur de l'écran ne permettrait sûrement pas de mettre que deux bandes.

8.3.3 Opportunité : le son

Le son n'a pas été utilisé pour l'application alors que c'est une des particularités des téléphones portables. Le son pourrait permettre de fournir plus d'informations aux utilisateurs. Différentes variables acoustiques comme le volume, le timbre et la position peuvent être utilisées pour représenter des aspects qualitatifs, quantitatifs et catégoriques de l'information. Ces variables sont aussi bénéfiques dans les circonstances où les représentations visuelles ne sont pas possible pour afficher des données ou bien pour enrichir des réalisations graphiques. Des graphiques utilisant le son ont déjà été réalisés au préalable tels qu'un Pie Chart [Franklin]. Le son aurait donc pu être utilisé pour le Pixel Bar Chart par exemple pour aider l'utilisateur à voir l'évolution d'un cluster, un son continu serait perçu par l'utilisateur, celui-ci augmentant si la valeur est haute et diminuant si celle-ci est basse. En général on utilise des tonalités pour exprimer différentes valeurs c'est à dire qu'une tonalité va représenter une valeur bien spécifique [Brown,Petrucci,Barass]. Mais Hermann [Hermann] a noté que les relations entre les sons et les données ne sont pas nécessairement intuitive.

Une autre idée serait d'enregistrer une voix prononçant chaque action (son verbal en fait). Dès lors l'utilisateur pourrait écouter le contenu de chaque cluster et ainsi on gagnerait de la place pour les visualisations. Cependant ceci est moins conseillé que d'écrire le contenu de chaque cluster car l'utilisateur aura plus difficile à retenir le contenu de ceux-ci si il les entend que si ils sont déjà écrits sur l'écran. Bien entendu, les deux peuvent être réalisés c'est à dire le contenu de chaque cluster écrit sur l'écran du GSM ainsi qu'une voix prononçant chaque action de ce cluster.

La navigation et les menus peuvent aussi être améliorés par le son comme mettre un son spécial pour chaque menu pour aider l'utilisateur à savoir où il est dans les différents menus hiérarchiques. Par exemple plus l'utilisateur va loin dans la navigation, plus le son est aigu, plus il est au début, plus le son est grave. Cependant, dans le contexte des téléphones portables, Helle [Helle] a montré que les utilisateurs coupent les sons des menus car ils les trouvent inappropriés.

Chapitre 9

Conclusions

A travers ce mémoire, notre but principal était de créer une application sur téléphones mobiles basée sur la combinaison de deux méthodes à savoir la classification non-supervisée et la visualisation. Cette application a pris comme exemple le domaine de la bourse mais ces deux techniques peuvent être utilisées dans tout autre domaine, où les données à afficher sont nombreuses.

Nous avons d'abord présenté les différents types de visualisations avant d'analyser celles se trouvant sur sites Wap boursiers. Une application existante, Midcast, a aussi été parcourue. Nous en avons conclu que sur la plupart des sites Wap boursiers, les visualisations d'information étaient peu lisibles. Aussi, les applications comme Midcast sont très intéressantes mais peu d'applications comme cela existent pour les téléphones portables.

Après, une analyse des besoins a été nécessaire pour distinguer les principaux types d'utilisateurs de l'application ainsi que leurs caractéristiques. Nous avons aussi présenté l'environnement de l'application de même que les différentes tâches que les deux types d'utilisateurs voudraient accomplir grâce à elle.

Puis, créer une application sur téléphones mobiles nécessite également de connaître les différentes limites de ceux-ci. La vitesse du processeur, la taille de la mémoire, la taille de l'écran, la navigation sont tant de caractéristiques limitées des GSM. Nous en avons conclu que pour programmer sur GSM, il faut savoir contourner ses différentes contraintes car certaines poseront toujours problèmes dans cette catégorie de "devices" (la taille de l'écran pour ne citer qu'elle).

Comme la classification non-supervisée (le clustering) est une des deux méthodes utilisées par notre application, nous avons donc décidé de l'analyser en détails et plus spécialement la technique des K-Means qui est la technique de classification non-supervisée que

nous avons choisie. Nous avons aussi vu que la technique du clustering peut présenter des avantages dans la visualisation d'information sur téléphones mobiles. Elle permet de résumer l'information en ce qu'elle a de plus pertinent et peut donc au départ tenir compte d'un plus grand nombre de données. Cette information réduite peut alors être affichée sur un petit écran.

Il restait donc ensuite à afficher sur l'écran du téléphone portable les informations boursières traitées par le clustering et nous avons pour cela utilisé quatre visualisations d'informations. Deux d'entre elles étaient des visualisations de base (le Bar Chart et le Tableau) qui offrent peu d'informations à l'utilisateur mais qui sont faciles à comprendre. Les deux autres, quant à elles, fournissent plus d'information mais nécessitent un apprentissage préalable de la part de l'utilisateur.

Pour réaliser cette application, nous avons choisi d'utiliser J2ME, une architecture technique de langage Java pour les systèmes embarqués. Dès lors nous avons pu analyser en détails le calcul des K-Means sur le serveur et le téléchargement des données boursières sur le Net. De même, l'interface de l'application fut aussi présentée. Plusieurs problèmes de l'application (comme la lenteur du téléchargement des données) ont été cités par la suite et finalement des solutions à ces problèmes ont été proposées.

Des tests utilisateurs ont été aussi effectué sur notre application sur une vingtaine de sujets. Cependant, les questions du tests étant très simples, aucune mesure de temps n'a été effectuée. Malgré tout, nous avons pu en ressortir que les utilisateurs ont des préférences pour certaines visualisations comme par exemple la Spirale qui n'a pas été bien accueillie en particulier.

Nous terminerons ce travail en mentionnant que cette application peut être améliorée de beaucoup. En effet, certains problèmes ont été perçus par les utilisateurs comme par exemple la navigation qui connaît quelques faiblesses. Comme nous l'avons également dit précédemment, le son pourrait être une opportunité pour améliorer les différentes visualisations. Pour finir, nous pouvons aussi indiquer que tous les utilisateurs ont trouvés l'utilisation de la classification non-supervisée ou clustering très intéressante particulièrement pour les téléphones portables.

Bibliographie

- [Ankerst] Ankerst M., Berchtold S., Keim D., Similarity clustering of dimensions for an enhanced visualisation of multidimensional data, presented at IEEE Symposium on Information Visualization, InfoVis'98, 1998.
- [Barass] Barass S., Zehner B., Responsive Sonification of Well-logs, Proceedings of the International Conference on Auditory Display ICAD 2000, Atlanta, April 2000.
- [Brown] Brown M.H, Herschberger J., Colour and sound in Algorithm Animation, IEEE Computer, Volume 25, number 12, pages 52-63, 1992.
- [Everitt] Everitt B.S., Cluster Analysis, Arnold, 1993.
- [Hardy] Hardy A., André P., An Investigation of Nine Procedures for Detecting the Structure in a Data Set, In Rizzi A., Vichi M. and Bock H.H. (eds), Advances in Data Science and Classification, Springer, 1998, pp. 29-36.
- [Helle] Helle S., Lepître G., Marila J., Laine P., Menu sonification in a mobile phone - A prototype study, Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1, 2001.
- [Hermann] Hermann T., Ritter H., Listen to your Data : Model-Based Sonification for Data Analysis, In M. R. Syed, Advances in intelligent computing and multimedia systems, Int. for Advanced Studies in System Research and Cybernetics, 1999.
- [Keim] Keim D., Hao Ming C. Ladisch J., Hsu M., Daya U., A New Technique for Visualizing Large Multi-Attribute Data Sets without Aggregation, Information Visualization Journal, Information Visualization Journal, 2002.
- [Keith] Franklin K., Roberts J.C., Pie Chart Sonification, Ebad Banissi et al, Proceedings Information Visualization (IV03), page 4-9, IEEE Computer Society, July 2003.
- [Mouchart] Mouchart M., Simar L., Recyclage en Statistique : Méthodes non-paramétriques, Comité de Statistique, Université Catholique de Louvain, 1978.
- [Noirhomme] Noirhomme M., De Baenst A., Check-list d'identification du contexte d'utilisation, Cours IHM de 1^{re} Maîtrise, FUNDP Namur, 2003-2004.
- [Petrucchi] Petrucci L.S., Harth E., Roth P., Assimacopoulos A., Pun T., Websound : a Generic Web Sonification Tool, and its Application to an Auditory Web Browser for Blind and Visually Impaired Users, ICAD 2000, Atlanta, April, 2000.

- [Spence] Spence R., Information Visualization, Addison Wesley, 2000.
- [S-Plus] Insightful's S-PLUS for Windows Homepage. Available from :
<http://www.insightful.com/products/product.asp?PID=3>
- [Weber] Weber M., Alexa M., Muller W., Visualizing Time-Series on Spirals, Proceedings of the IEEE Symposium on Information Visualization (InfoVis), IEEE Press, 2001, pp. 7-14.
- [Wonnacott & Wonnacott91] Thomas H. Wonnacott, Ronald J. Wonnacott, "Statistique : Economie - Gestion - Sciences - Médecine (avec exercices d'application)", 4ème Edition, Economica, Paris, France, 1991.

Annexe A

Appendice

A.1 Les différents questionnaires

A.1.1 Utilisateur 1

Profession	Etudiant
Domaine de la profession	Informatique
Age	21
Sexe	Feminin
Expérience dans le domaine de l'informatique	4
Expérience dans les domaines liés à la bourse	0
Expérience dans les GSMs	2

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Pixel Bar Chart un peu lent.
- Problème avec le bouton Back.
- Pour l'intitulé des boutons, tout mettre en Français ou en Anglais.
- Difficulté de visualiser les informations pour le Pixel Bar Chart.
- Plus facile d'utiliser le Bar Chart de type 2 que de type 1 quand il y a des positifs et des négatifs.
- Clustering est une bonne technique pour les mobiles.

A.1.2 Utilisateur 2

Profession	Professeur
Domaine de la profession	Managment
Age	57 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	34 ans
Expérience dans les domaines liés à la bourse	30 ans
Expérience dans les GSMs	7 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Apprentissage rapide.
- Problème avec le bouton Select/Ok.
- Mettre l'application sur une évolution journalière.
- Clustering ++.

A.1.3 Utilisateur 3

Profession	Etudiant
Domaine de la profession	Informatique
Age	23 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	6 ans
Expérience dans les domaines liés à la bourse	1 ans
Expérience dans les GSMs	3 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Problème avec le Select/Ok.
- Spirale un peu compliquée à lire.
- Pixel Bar Chart : pas assez de couleurs sur le GSM pour voir l'évolution.
- Le reste est bien.

A.1.4 Utilisateur 4

Profession	Etudiant
Domaine de la profession	Informatique
Age	21
Sexe	Masculin
Expérience dans le domaine de l'informatique	4
Expérience dans les domaines liés à la bourse	2
Expérience dans les GSMs	2

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Combo Box à vérifier.
- Seule difficulté : variations des couleurs.
- Manipulation moins intuitive que sur un PDA.

A.1.5 Utilisateur 5

Profession	Etudiant
Domaine de la profession	Informatique
Age	23 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	10 ans
Expérience dans les domaines liés à la bourse	6 mois
Expérience dans les GSMs	5 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Problème de bouton Select/Ok.
- Pixel Bar Chart et Spirale trop lents.
- Spirale difficile à lire.
- Visualisation rapide. Clustering : plus facile à lire que si tous les chiffres étaient écrits (vu le nombre de valeurs par cluster).

A.1.6 Utilisateur 6

Profession	Etudiant
Domaine de la profession	Informatique
Age	22 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	13 ans
Expérience dans les domaines liés à la bourse	3 mois
Expérience dans les GSMs	3 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Navigation : Select/Ok et bouton Back mauvais.
- Ce type d'application n'a pas lieu d'être sur un GSM.
- Les visualisations se valent toutes mais le Bar Chart de type 2 est bof bof.

A.1.7 Utilisateur 7

Profession	Gérant
Domaine de la profession	Station essence
Age	34 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	1 ans
Expérience dans les domaines liés à la bourse	0
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Il faudrait un GSM plus rapide.
- Le clustering est parfait pour les GSM.

A.1.8 Utilisateur 8

Profession	Etudiant
Domaine de la profession	Informatique
Age	24 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	8 ans
Expérience dans les domaines liés à la bourse	1 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Il faudrait une échelle de couleurs pour la Spirale.
- Le Tableau est déjà intégré dans les autres visualisations.
- Le clustering peut aider pour les séries temporelles.
- Problème de la sélection d'éléments : bouton Select/Ok.

A.1.9 Utilisateur 9

Profession	Etudiant
Domaine de la profession	Informatique
Age	24 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	4 ans
Expérience dans les domaines liés à la bourse	6 mois
Expérience dans les GSMs	2 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Navigation est plus ou moins intuitive, il y a moyen de l'améliorer.
- Améliorer le bouton Select/Ok.
- Proposer des couleurs différentes.
- Faire une légende.

A.1.10 Utilisateur 10

Profession	Etudiant
Domaine de la profession	Informatique
Age	23 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	20 ans
Expérience dans les domaines liés à la bourse	2 ans
Expérience dans les GSMs	3 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- C'est lent.
- Le backlight s'éteint trop vite.
- Pourquoi Option->Select puis Option->Ok au lieu de Select et Ok ?
- Pas de retour au début directement.
- Les couleurs sont biens.
- Ecran trop petit.
- Clustering est bien.

A.1.11 Utilisateur 11

Profession	Etudiant
Domaine de la profession	Ingénieur Civil
Age	23 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	5 ans
Expérience dans les domaines liés à la bourse	2 ans
Expérience dans les GSMs	5 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Bonnes visualisations : Bar Chart et Tableau.
- Mauvaise visualisation : Spirale car manque de clarté au niveau des couleurs. Une visualisation linéaire serait beaucoup plus claire que celle de la Spirale.
- Clustering : je n'avais jamais utilisé des technologies de ce type sur GSM. Mon avis est très positif. Beaucoup de domaines d'applications possibles.

- Selection des options dans le menu des choix des périodes et valeurs pas très claire.
- Il faudrait plus d’ergonomie dans le choix des options.

A.1.12 Utilisateur 12

Profession	Etudiant
Domaine de la profession	Informatique
Age	24 ans
Sexe	Masculin
Expérience dans le domaine de l’informatique	17 ans
Expérience dans les domaines liés à la bourse	1 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Certains textes sont rognés par le menu inférieur.
- Bar Charts différentiels intéressant.
- Clustering intéressant.
- Les Bar Charts sont les plus efficaces.
- Je préférerais un graphique pour l’évolution.
- Encadrer le cluster actif d’une autre couleur.
- Bouton Back qui disparaît dans certains cas.
- Save non automatique lors de sélections.

A.1.13 Utilisateur 13

Profession	Etudiant
Domaine de la profession	Ingénieur civil
Age	20 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	1 ans
Expérience dans les domaines liés à la bourse	1 ans
Expérience dans les GSMs	4 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Lent pour le Pixel Bar Chart.
- Les mots sont coupés à la fin dans le contenu des clusters.
- Ce n'est pas toujours le même bouton retour pour les menus.
- La couleur bleu dans le Bar Chart ne permet pas de voir l'indicateur clignotant correctement.
- Facile à utiliser.
- Clustering : c'est intéressant de regrouper les informations entre elles sur GSM.
- Navigation simple mais il faudrait créer un accès plus rapide au menu de départ.
- Avoir le choix des couleurs.

A.1.14 Utilisateur 14

Profession	Employé
Domaine de la profession	Culture
Age	27 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	5 ans
Expérience dans les domaines liés à la bourse	10 ans
Expérience dans les GSMs	10 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Pixel Bar Chart beaucoup trop lent.
- Spirale pas assez précise.
- L'application est très pratique pour être informé et très précis suivant le type de graphique.

A.1.15 Utilisateur 15

Profession	Employé
Domaine de la profession	Graphisme
Age	23 ans
Sexe	Féminin
Expérience dans le domaine de l'informatique	6 ans
Expérience dans les domaines liés à la bourse	6 mois
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Pixel Bar Chart est trop lent.
- La Spirale n'est pas assez claire.
- Certaines visualisations permettent d'avoir une visualisation claire de l'évolution des clusters, par contre d'autres ne sont pas évidentes.
- Une fois le principe de la navigation compris, celle-ci est intuitive.
- Ralentissement.

A.1.16 Utilisateur 16

Profession	Etudiant
Domaine de la profession	Comptabilité
Age	20 ans
Sexe	Féminin
Expérience dans le domaine de l'informatique	1.5 ans
Expérience dans les domaines liés à la bourse	3 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- La Spirale est très compliquée à lire.
- Bouton Select/Ok pas très facile à utiliser.
- Le clustering est pratique pour les GSM à cause de la taille de l'écran.
- Une échelle de couleurs pour la Spirale et le Pixel Bar Chart serait bien.

A.1.17 Utilisateur 17

Profession	Professeur
Domaine de la profession	Informatique de Gestion
Age	32 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	20 ans
Expérience dans les domaines liés à la bourse	8 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Pixel Bar Chart est super.
- Bar Chart de type 2 un peu perturbant.
- Lisibilité est bonne.
- Valeur : attention à la longueur de la période. Proposer une période très courte.
Mais une longue période est bien pour le Pixel Bar Chart.
- Étendre l'application à la prédiction.
- Navigation pas uniforme.

A.1.18 Utilisateur 18

Profession	Gérante
Domaine de la profession	Vente
Age	53 ans
Sexe	Féminin
Expérience dans le domaine de l'informatique	1 ans
Expérience dans les domaines liés à la bourse	5 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- La Spirale est trop compliqué pour comparer deux clusters.
- Le Pixel Bar Chart est bien pour voir l'évolution malgré qu'il est lent.
- Le Bar Chart de type 1 est mieux que celui de type 2.
- Navigation un peu perturbante, surtout pour le Select/Ok.

A.1.19 Utilisateur 19

Profession	Etudiant
Domaine de la profession	Comptabilité
Age	23 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	5 ans
Expérience dans les domaines liés à la bourse	5 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- Bar Chart de type 2 est plus compliqué que celui de type 1.
- Bouton Back n'est pas bien.
- Choisir des couleurs différentes pourrait être bien.
- Le clustering est pratique pour les GSM et doit pouvoir servir à l'avenir dans ce domaine.
- Lisibilité est parfois dure pour le Pixel Bar Chart et la Spirale.

A.1.20 Utilisateur 20

Profession	Etudiant
Domaine de la profession	Informatique
Age	22 ans
Sexe	Masculin
Expérience dans le domaine de l'informatique	7 ans
Expérience dans les domaines liés à la bourse	3 ans
Expérience dans les GSMs	6 ans

Bar Chart 1	Mauvais	Moyen	Bon	Excellent
Bar Chart 2	Mauvais	Moyen	Bon	Excellent
Tableau	Mauvais	Moyen	Bon	Excellent
Pixel Bar Chart	Mauvais	Moyen	Bon	Excellent
Spirale	Mauvais	Moyen	Bon	Excellent

Commentaires

- La dernière ligne dans le contenu des clusters est coupée.
- Le Pixel Bar Chart est très intéressant car la taille de l'écran est petite et c'est assez intelligent d'utiliser les couleurs pour représenter les informations.
- Problème avec le bouton Select/Ok.
- Un bouton de retour au menu de départ serait bien.

A.2 Table des valeurs critiques pour la t distribution de Student

	$\alpha = 0,50$	$\alpha = 0,20$	$\alpha = 0,10$	$\alpha = 0,050$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,002$	$\alpha = 0,001$
<i>ddl</i>	$t_{0,25}$	$t_{0,10}$	$t_{0,05}$	$t_{0,025}$	$t_{0,01}$	$t_{0,005}$	$t_{0,001}$	$t_{0,0005}$
1	1,000	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,816	1,886	2,920	4,303	6,695	9,925	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,684	1,316	1,708	2,060	2,405	2,787	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421

	$\alpha = 0,50$	$\alpha = 0,20$	$\alpha = 0,10$	$\alpha = 0,050$	$\alpha = 0,02$	$\alpha = 0,01$	$\alpha = 0,002$	$\alpha = 0,001$
ddl	$t_{0,25}$	$t_{0,10}$	$t_{0,05}$	$t_{0,025}$	$t_{0,01}$	$t_{0,005}$	$t_{0,001}$	$t_{0,0005}$
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307	3,551
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232	3,460
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160	3,373
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090	3,291
	$Z.25$	$Z.10$	$Z.05$	$Z.025$	$Z.010$	$Z.005$	$Z.0010$	$Z.0005$